

T.M.  
(043)53  
2021  
Ab14

## TESIS DE MAESTRÍA EN CIENCIAS FÍSICAS

### MODELOS DE SELECCIÓN NATURAL BASADOS EN ENERGÍA

**Noam Abadi**  
Maestrando

**Dr. Guillermo Abramson**  
Director

#### **Miembros del Jurado**

Dr. Sebastián Bouzat  
Dr. Marcelo Kuperman  
Dra. Karina Laneri

Febrero de 2021

División Física Estadística e Interdisciplinaria – Centro Atómico Bariloche

Instituto Balseiro  
Universidad Nacional de Cuyo  
Comisión Nacional de Energía Atómica  
Argentina

INVENTARIO: 24136

05.07.2021

Biblioteca Leo Falicov

A mi viejo, que me enseñó a apuntar alto.  
A mi vieja, que me enseñó que no está mal no hacerlo.  
A mi hermano, que me recuerda lo que es la libertad.  
A mi hermana, que me recuerda lo que es querer.  
A Seba, con quien desarmé todo lo que se nos cruzó.  
A Marcos, Jose, Oscar, Sergio y Mauri, que me hicieron más compañía de lo que se  
imaginan.  
A Bauti y Nahue, quienes me ayudaron a lidiar con mi procrastinación compulsiva.  
A Mati, Pablo y Bruno, que me hacen reir muchísimo.  
Finalmente, a Taïs, que está en todas las anteriores y más.



# Índice de contenidos

|  |     |
|--|-----|
| Índice de contenidos                     | v   |
| Resumen                                  | vii |
| Abstract                                 | ix  |
| 1. Introducción                          | 1   |
| 2. Modelo de redes de Kauffman           | 7   |
| 2.1. Un ejemplo simple                   | 8   |
| 2.2. Limitación de recurso               | 11  |
| 2.3. Errores genéticos                   | 15  |
| 2.4. Herencia                            | 17  |
| 3. Modelo de información                 | 21  |
| 3.1. Transmisión vertical de información | 21  |
| 3.2. Información interna                 | 31  |
| 4. Modelo de la hidra inmortal           | 33  |
| 4.1. Dinámica poblacional                | 35  |
| 4.2. Más allá de los modelos estándar    | 38  |
| 4.3. Dinámica genética                   | 41  |
| 5. Modelo de la hidra                    | 45  |
| 5.1. Dinámica poblacional                | 45  |
| 5.2. Dinámica individual                 | 46  |
| 5.2.1. Genética                          | 46  |
| 5.2.2. Probabilidad de muerte            | 50  |
| 5.2.3. Comportamiento altruista          | 52  |
| 5.2.4. Campo medio                       | 53  |
| 5.3. Dinámica del recurso                | 55  |
| 5.4. Algunas modificaciones              | 58  |
| 6. Discusión y conclusiones              | 61  |
| A. Pseudocódigo                          | 65  |
| Bibliografía                             | 67  |
| Publicaciones asociadas                  | 71  |

|                 |
|-----------------|
| Agradecimientos |
|-----------------|

|    |
|----|
| 73 |
|----|

# Resumen

Desde el trabajo de Darwin, *El origen de las especies*, la selección natural (supervivencia de los más aptos) ha sido cada vez más aceptado como el mecanismo por el cual las especies evolucionan y se extinguen. A pesar de su creciente popularidad, no hay un consenso general acerca de cómo determinar cuáles son, realmente, las especies más aptas.

Para solucionar esto, proponemos una manera de estudiar la selección natural que no se basa en buscar una función a optimizar sino en imponer una restricción fundada en las necesidades energéticas de los seres vivos. Nuestro modelo captura la forma general en la que las características fisiológicas (fenotipo) codificadas en el ADN (genotipo) determinan el balance entre sus costos y retribuciones. De generación en generación, los parámetros de esas restricciones se mutan aleatoriamente, y los individuos no tienen ningún objetivo en particular. Ni siquiera el de sobrevivir está impuesto como un objetivo en sí, sino que lo hacen los que pueden. Al incorporar una competencia por una fuente de energía compartida (recurso) surge la noción de aptitud genética en la supervivencia reiterada de individuos con ciertos rasgos comunes.

Estudiamos algunos casos particulares de la dinámica energética que surge del balance entre costo y retribución. En primera instancia, cadenas booleanas cuyos bits interpretamos como genes activos o inactivos (y asociados con un consumo, retribución máxima, y un estado booleano sucesivo de cada individuo) sirvieron para estudiar distintas dinámicas de recuperación de recurso en “seres vivos” más similares a moléculas prebióticas. Luego, para estudiar la transmisión de información genética, nos deshicimos de los estados booleanos y tratamos con un consumo, retribución máxima y tasa de reproducción constantes para cada individuo, permitiendo definir una “distancia” genética de manera más objetiva. En una tercera etapa, consideramos el costo (principalmente metabólico y cinético) que tiene asociado un cuerpo de cierto tamaño, pudiendo moverse a una velocidad limitada por el mismo. Finalmente, agregamos un grado de libertad al modelo anterior, con la velocidad independiente al tamaño, para producir un escenario más realista. Hicimos variaciones a los diferentes casos y observamos cómo estas influían en los resultados ya reportados.

Hallamos que muchos de los resultados conocidos de la biología teórica y cuantitativa surgen en este tipo de modelos sin incorporarlos explícitamente. Estos incluyen pero no se limitan a la exclusión competitiva, el equilibrio puntuado, la sensibilidad a condiciones iniciales, la convergencia evolutiva, la especiación, y el altruismo grupal a pesar de reglas egoístas individuales.

**Palabras clave:** BIOFÍSICA, EVOLUCIÓN, FITNESS, SELECCIÓN NATURAL, ESTADÍSTICA FUERA DEL EQUILIBRIO.





# Abstract

Ever since Darwin's *The origin of species*, natural selection (survival of the fittest) has been more and more accepted as the pathway through which species evolve and go extinct. Despite its growing popularity, there is no general consensus in determining which are, in fact, the fittest species.

To face this problem, we propose a way to study natural selection without imposing a function to be optimized, but a restriction given by energetic necessities of living beings. Our model captures the way in which physiological traits (the phenotype) encoded by the DNA (genotype) determine the balance between costs and retributions of said traits. From one generation to the next, the parameters involved in the restrictions mutate randomly, with no general goal implied for individuals. Not even survival is imposed as a goal, since the ones that do so are merely the ones that can. Through competition for a shared energy source (resource), a notion of fitness emerges in the repeated survival of individuals with certain common traits.

We study some special cases of the energy dynamics that result from the balance between cost and retribution. In a first scenario, we use boolean chains whose bits we interpret as active or inactive genes (associated with an energy consumption, maximum retribution and following boolean state for each individual) to study different forms of resource recovery in "living beings" more similar to prebiotic molecules. Next, we drop the boolean states and focus only on consumption, retribution and reproduction frequency to get a clearer idea of the relation between genetic and phenotypic similarity, and the influence of transmitting genetic information. In a third stage, we devise a model for the metabolic and kinetic cost of an organism of a certain body size, with an associated velocity limited by said size. Finally, we add a degree of freedom to the former model, uncoupling velocity from body size to produce a more realistic scenario. We also vary the different cases, observing how these changes influence reported results.

With this method, we have found that many of the known results of theoretical and quantitative biology emerge without their explicit inception in the model. These include (but are not limited to) competitive exclusion, punctuated equilibrium, sensitivity to initial conditions, evolutionary convergence, speciation and group altruism from selfish individual rules.

**Keywords:** BIOPHYSICS, EVOLUTION, FITNESS, NATURAL SELECTION, OUT OF EQUILIBRIUM STATISTICS.







# Capítulo 1

## Introducción

*“...and all the science I don’t understand,  
It’s just my job five days a week...”*  
— Elton John

En el principio, el universo estaba en caos. Luego, desde el momento en el que se pudo hablar de momentos, Poderes omnipotentes, omnipresentes e inmortales comenzaron a remover el caos. La nube de materia y energía comenzó a tomar forma. Pequeñas fluctuaciones en su distribución, inicialmente tan uniforme, comenzaron a formar grumos. Los grumos, casi caprichosamente, se atraían o repelían entre sí, formando grumos mayores. Estos se comportaban según las reglas impuestas por los Poderes inmortales, pero la forma en la que esto se evidenciaba era distinta cada vez que se agregaba un nivel de complejidad. Algunos momentos después del principio<sup>1</sup>, se hizo la luz. Algunos grumos se habían reunido tan densamente que, al chocar, se unían liberando energía y logrando una nueva estabilidad. Otros grumos, en aglomeraciones menos densas, podían aprovechar esta energía para lograr combinaciones antes inaccesibles. Esta es la historia de estos grumos, de cómo se descubrieron a sí mismos, y principalmente de cómo, al darse cuenta que no eran tan distintos al resto del universo, comenzaron a preguntarse cómo habían llegado allí y hacia dónde iban.

Por supuesto, no podemos ni pretendemos descubrir y contar la historia completa de cómo llegamos hasta aquí porque no existe una manera objetiva de hacerlo. Siempre se hacen suposiciones para dar una explicación y estas podrán generar más preguntas hasta que se haya llegado a un nivel de hipótesis comunes con la persona que las hace. Sin embargo, para otra persona, posiblemente estas suposiciones también requieran explicación. Partiremos, entonces, desde la base de la biología. Daremos una breve reseña histórica de la evolución de esta ciencia y sus aportes a las preguntas *¿De dónde venimos?* y *¿Hacia dónde vamos?* Luego, daremos un resumen de los problemas que se están intentando responder actualmente y cómo este trabajo se enmarca, complementa y diferencia de ellos.

La biología es, discutiblemente, una de las ciencias más antiguas de mundo. A pesar de que no estaba asentada como una ciencia formal, sociedades de cazadores-recolectores debían tener conocimientos, transmitidos de generación en generación, de lo que actualmente se llama zoología y botánica [1] para saber qué animales y plantas (y qué partes de ellos) consumir. Los primeros registros son de sociedades asirias y babilónicas, con esculturas de bajo relieve que muestran conocimientos de medicina veterinaria y la reproducción sexual de la palmera datilera desde hace entre 5500 y 3800 años [2].

Las primeras teorías sobre los orígenes y la fuerza motora de la vida (más allá de nociones religiosas)

---

<sup>1</sup>Un conjunto particular de grumos (sobre un grumo mucho mayor que orbita a un grumo aún más grande, ninguno de los cuales existía en ese entonces) diría que fueron muchos, no algunos.

pueden atribuirse a Tales de Mileto<sup>2</sup>. Sostenía que todos los seres vivos están compuestos de agua, y que el universo contenía una fuerza creadora de vida llamada *physis* [3]. Anaximandro, un estudiante de Tales, propuso que la vida requería también de tierra y *apeiron*, una sustancia que hoy llamaríamos gaseosa. También presentó la idea de que la vida emergió del barro y evolucionó por transmutación en lo que conocemos hoy en día citeAnaximanderFragments. La teoría de que la vida se originó en barro sigue siendo estudiada actualmente [4].

Una de las primeras clasificaciones taxonómicas<sup>3</sup> de los seres vivos se le atribuye, por supuesto, a Aristóteles<sup>4</sup>. En su trabajo *Historia de los animales* [5], hace una separación por características objetivas (principalmente con o sin sangre y número de patas) pero también algunas más subjetivas, de acuerdo a lo que él llamaba alma (*psyche*), que le permite una clasificación de todos los seres en racionales, sensitivos o vegetativos. Luego de Aristóteles, la ciencia vio poco progreso durante aproximadamente dos milenios. La cultura islámica fue la que se encargó de traducir, estudiar y transmitir las obras de la Antigüedad clásica a fines de la Edad Media [6]. Durante los siglos XII y XIII, los escritos del teólogo persa Al-Ghazali<sup>5</sup> cobraron influencia, promoviendo la disección como método para ganar conocimiento sobre la creación de Allah [7].

Desde mediados del siglo XVI en Europa, con el Renacimiento, comenzó una serie de cambios en la manera de estudiar y registrar el mundo natural. La medicina comenzó a adquirir conocimientos de manera más rigurosa con la disección de cadáveres humanos [8]. Los avances pasaron de basarse en la escuela Escolástica<sup>6</sup> del Medioevo al Empiricismo de la revolución científica [9, 10]. Artistas del Renacimiento, interesados en mejorar la calidad de sus obras, colaboraban con naturalistas, impulsando los avances del conocimiento científico [11]. Leonardo da Vinci<sup>7</sup> destacó las diferencias y similitudes entre la anatomía humana y de otros animales [12], dando los primeros pasos que conducirían a descubrir que los seres humanos no somos sino una especie más del reino animal. El perfeccionamiento del microscopio por Antonie van Leeuwenhoek en 1674 permitió observar, por primera vez, una célula viva<sup>8</sup> [13]. En 1839, Schwann y Schleiden propusieron una teoría en la que conjeturaban que la célula era la unidad fundamental de la vida, cerrando aun más la brecha entre los seres humanos y otros animales [14].

El proceso de unificación fue avanzado y prácticamente culminado por Jean-Baptiste Lamarck<sup>9</sup> y Charles Darwin<sup>10</sup>, cada uno con su propia teoría para explicar cómo las especies (incluyendo la humana) interactuaban y daban lugar a especies nuevas. Para Lamarck, la acción ejercida por el entorno sobre los individuos los hacía variar levemente en sus características (como una jirafa estirando su cuello para llegar a las hojas más altas), y estas características adquiridas se heredaban a las generaciones siguientes. Para Darwin, en cambio, los individuos nacían con un conjunto de características que se transmitía con pequeños cambios de generación en generación (aunque murió sin saber cual era el mecanismo de transmisión). Estas les otorgaban ciertas ventajas (o desventajas) respecto a otros individuos con los que competían. A su vez, esto generaba una presión selectiva sobre qué individuos se reproducían más exitosamente (los “más aptos” de su subtítulo), dejando más individuos similares, pero también levemente distintos. Su teoría, que resume en su libro *Sobre el origen de las especies* [15] asienta no solo la fundación de la teoría evolutiva moderna sino también las bases de este trabajo.

<sup>2</sup>Mileto, antigua Grecia, hace aproximadamente 2700 años. Es el autor del Teorema de Tales sobre la proporcionalidad de los segmentos que se obtienen cortando paralelas con líneas rectas. Inspiró una canción del grupo musical Les Luthiers.

<sup>3</sup>En función de características comunes, generalmente fisiológicas.

<sup>4</sup>Estagira, antigua Grecia, hace aproximadamente 2400 años. No hay ciencia moderna en la que no haya dejado su marca.

<sup>5</sup>Tus, Imperio Seljuk —actualmente Irán— hace aproximadamente mil años.

<sup>6</sup>Basada en la argumentación y el renombre.

<sup>7</sup>Vinci, Florencia —actualmente Italia— siglos XV a XVI. Fue el Aristóteles del Renacimiento.

<sup>8</sup>Aunque Robert Hooke (Isla de Wight, Inglaterra, 1635–1703) fue el primero en ver células (muertas) en corcho en 1665, y quien les dio el nombre.

<sup>9</sup>Bazentin, Francia, 1744–1829.

<sup>10</sup>Shrewsbury, Inglaterra, 1809–1882.

A principios del siglo XX, el descubrimiento de la difracción de rayos X de von Laue<sup>11</sup> [16] y el redescubrimiento de las leyes de Mendel hicieron prosperar el conocimiento sobre los mecanismos de transmisión de la información genética [17]. Para mediados del siglo XX los cromosomas ya se entendían lo suficientemente bien como para asentar la teoría de los mecanismos de la herencia en una base sólida. El ácido nucleico y sus cinco bases, que habían sido aislados por Kossel en 1878 ya se conocían, dejando listos los elementos necesarios para el primer modelo correcto de doble hélice de Jim Watson, Francis Crick y Rosalind Franklin de 1953. La biología cuantitativa también tomó impulso en este siglo, comenzando con los trabajos de Alfred Lotka y Vito Volterra<sup>12</sup> sobre la dinámica poblacional de los sistemas de depredadores y presas [20]. A lo largo del siglo, esta rama de la biología se vio impulsada por el aumento de la cantidad de los datos disponibles, el desarrollo de la teoría del caos, y el incremento del poder computacional [21, 22].

Curiosamente, la biología ha pasado toda su historia estudiando la vida sin saber lo que la vida es [23]. Muchas definiciones han sido propuestas [24], pero ninguna es totalmente satisfactoria. Hasta podría decirse que la línea divisoria entre los seres vivos y los objetos inanimados se ha vuelto más difusa con el desarrollo de modelos computacionales crecientemente complejos, la inteligencia artificial o incluso los virus de computadoras [25]. En la rama más filosófica de esta disciplina, características que Aristóteles atribuía al alma, como la inteligencia, la conciencia y la capacidad de aprendizaje, tampoco están bien definidas.

En el área de la genética, la biología está avanzando a tasas sin precedentes. Los niveles de precisión permiten en la actualidad hallar secuencias específicas asociadas con enfermedades genéticas en el ADN, y se están desarrollando herramientas como CRISPR [26, 27] que permiten “editar” genes individuales, potencialmente erradicando enfermedades pero también planteando grandes dilemas éticos. El año pasado (2020) y este han sido fuertemente marcados por la pandemia del virus SARS-CoV-2, y la secuenciación en tiempo real de su código genético [28], combinada con conocimientos nuevos y viejos, modelos teóricos y empíricos y el trabajo conjunto de la población nos han permitido salir adelante.

La teoría evolutiva, por otra parte, ha encontrado aplicaciones inesperadas. Métodos de optimización utilizan algoritmos inspirados en la selección natural [29]. Los aportes del equilibrio puntuado [30], sistemas y redes complejas [31] y el desarrollo de la termodinámica fuera del equilibrio [32], la han asentado sobre bases teóricas más rigurosas. Una de las preguntas abiertas de esta ciencia es el objetivo de la optimización evolutiva [33]. Esto tiene el potencial para responder a la pregunta *¿A dónde vamos?* a la que nos referimos al principio de este capítulo. A pesar de muchas propuestas interesantes [34–36], principalmente basadas en el desarrollo de los procesos termodinámicos fuera del equilibrio [32], no hay un consenso generalizado. Es en este contexto que se enmarca esta tesis.

Comenzamos apuntando a responder una pregunta muy puntual: ¿A qué se refería Darwin con *aptitud* cuando dijo que sobreviven los más aptos? Desde ya adelantamos que no logramos responder a la pregunta de manera tan general como está planteada. Sin embargo, logramos proponer una forma robusta de encararla en distintos escenarios.

Esta se logra en dos partes. Primero, olvidemos que estamos buscando a los *más* aptos. Los *más* aptos sobreviven como tales solo cuando deben competir con otros. Una forma más honesta de postular el principio de selección natural es que sobreviven los que pueden. Esto transforma la manera habitual de entender la selección natural como la optimización de una *fitness* reproductiva, en una restricción. Naturalmente, surge la pregunta ¿Cuáles son los que *pueden*? Para responderla recordamos uno de los

<sup>11</sup>Pfaffendorf, Imperio Alemán, 1879–1960.

<sup>12</sup>Aunque los trabajos de Thomas Malthus (Westcott, Inglaterra, 1766–1834) [18], y Pierre Verhulst (Bélgica, 1804–1849) [19], preceden a sus trabajos por más de un siglo.

principios fundamentales de la física, el de la conservación de la energía. De acuerdo a este, la energía  $E$  que necesitan los seres vivos para realizar sus funciones vitales solo puede provenir de un recurso,  $f$ , que es parte del entorno donde viven los individuos. Conseguir este recurso tiene asociado un costo energético,  $C$ , y a su vez produce una retribución,  $\varphi$ . Luego del consumo, el recurso tendrá alguna dinámica de recuperación,  $R(f)$ , cuya fuente de energía no se detalla (y que en la Tierra es, en última instancia, el Sol). Podemos plantear entonces la siguiente dinámica para la energía y su fuente:

$$\frac{dE}{dt} = -C(\mathbf{F}) + \varphi(\mathbf{F}, f) \quad (1.1)$$

y

$$\frac{df}{dt} = R(f). \quad (1.2)$$

En la Ecuación (1.1) asociamos a cada fenotipo  $\mathbf{F}$  un costo energético, y a cada par recurso-fenotipo una retribución [31]. Así, se define un modelo para la dinámica energética de un individuo a partir de sus características fisiológicas y su entorno. A pesar de que las ecuaciones (1.1) y (1.2) puedan tomar formas muy simples, la dinámica del sistema completo puede ser bastante complicada por tratarse de un sistema de ecuaciones estocásticas acopladas de manera no lineal. Por esta razón, la mayoría de los resultados presentados en este trabajo fueron obtenidos mediante simulaciones numéricas. En el Apéndice A se presentan los algoritmos generales de actualización de la población y el recurso, común a todos los modelos.

Debido a que dijimos que la energía es la que los individuos necesitan para vivir, diremos también que cuando la energía llega a cero, el individuo muere. Destacamos que lo recíproco no necesariamente es cierto, ya que una muerte por envejecimiento, por ejemplo, no implica inanición.

Cabe destacar que la energía puede ser interpretada como una métrica para la calidad de vida de un individuo, dado su entorno. En muchas situaciones reales, esta queda determinada por una serie de propiedades en vez de por la energía exclusivamente. Por ejemplo, la mayoría de los seres vivos necesitan agua y nutrientes para que sus células puedan recibir la energía que necesitan. En este caso, una descripción macroscópica más adecuada se lograría reemplazando el escalar energía por un vector que, en sus componentes, indique la disponibilidad de alimento y agua. Obviamente, el consumo y la retribución también deberían reemplazarse acordemente con modelos apropiados para cada una de las componentes. Por simplicidad, en esta tesis estudiaremos modelos donde las propiedades son escalares.

Otro elemento fundamental de la vida, y por lo tanto de la selección natural, es la reproducción. En nuestro modelo, el genotipo  $\mathbf{G}$  de un adulto que se reproduce se transmite a la descendencia con algún grado de mutación y se recibe como  $\mathbf{G}'$ . El nuevo genotipo codifica las condiciones iniciales del nuevo fenotipo  $\mathbf{F}'(t=0)$  y su evolución temporal  $\mathbf{F}'(t)$ . El grado de mutación y la métrica de similitud entre progenitores y descendientes dependerá del modelo.

En definitiva, nos interesa estudiar las propiedades que tiene la selección natural bajo estas suposiciones generales. Para hacerlo propondremos algunos modelos particulares para las relaciones entre genotipos, fenotipos, consumos, retribuciones, reproducción y dinámica de recurso. Priorizaremos la simplicidad sobre el realismo en los modelos, pero intentaremos mantenernos dentro de lo verosímil. Empezaremos con un modelo muy simple que representa de manera bastante precisa el código genético presente en los organismos reales, pero sin mucha interpretación fenotípica. Mostraremos que incluso en este caso, tiende a sobrevivir una sola especie como indica el principio de exclusión competitiva [37]. Continuaremos simplificando este modelo para estudiar cómo influye la transferencia de información en este tipo de modelos y veremos cómo influye agregar grados de libertad por los cuales los individuos pueden evolucionar. Luego, nos centraremos en un modelo basado en organismos reales para analizar casos más complejos con dinámicas de recurso distintas. Compararemos la dinámica poblacional de



este modelo con las de otros conocidos para encontrar similitudes y diferencias. Terminaremos con una modificación del modelo anterior (que agrega un grado de libertad) y de la cual surgen una multitud de resultados, incluyendo especiación, altruismo a pesar de reglas egoístas y movimiento inteligente rudimentario.



## Capítulo 2

# Modelo de redes de Kauffman

*“From so simple a beginning, endless forms most beautiful have been, and are being, evolved.”*

— Nightwish

En este capítulo estudiaremos un modelo muy simple para comenzar a intuir el comportamiento de modelos que obedecen a las generalidades expuestas en el Capítulo 1. El modelo se basa en las ideas de redes booleanas aleatorias de Stuart Kauffman [38]. A su vez, las redes booleanas están inspiradas en un modelo de transcripción y traducción del ADN en proteínas, *El libro de Kauffman es como una recopilación de todos sus papers, me parece que esto puede quedar sin cita.* por lo cual tienen un valor conceptual importante.

La idea general es que las proteínas actúan como activadoras o inhibidoras de la expresión de los genes a partir de los cuales ellas mismas se sintetizan. Esto ocurre de manera cruzada y en general intrincada, de manera que el sistema de genes funciona como una compleja red de activación y represión, en la cual el estado de activación a un dado tiempo determina el estado a un tiempo posterior.

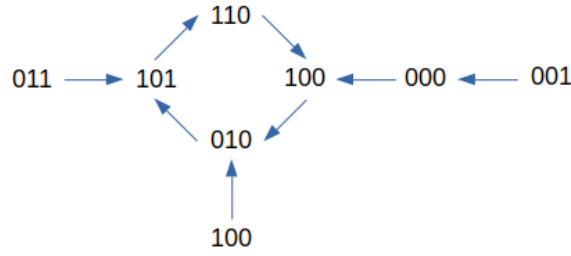
En este marco, los estados de una red de  $L$  genes se pueden representar por variables booleanas de  $L$  bits, donde un bit que toma el valor 1 significa que el gen está activo (expresándose) y un bit que toma el valor 0, que está inactivo (no expresado). Como no se le da una interpretación específica a los genes en este modelo (no hay un gen que determine, por ejemplo, el tamaño del individuo), los estados sucesivos son determinados de manera aleatoria, pero quedan fijos para cada realización. Para una cadena de ADN de  $L$  genes ( $2^L$  estados posibles) que comienza en un estado  $s_o = \mathbf{G}$ , los estados sucesivos por los que pasa un individuo están determinados por un mapeo como el que mostramos en las primeras dos columnas de la Tabla 2.1, ejemplificado en un sistema de  $L = 3$  genes. Los estados  $s_i$  representan los sucesivos fenotipos del individuo. En el modelo de Kauffman, estos estados se asocian a una *fitness*. Como el objetivo de este trabajo es que el *fitness* sea una propiedad emergente, asociamos cada uno de estos estados con un costo energético  $C$  y un pedido de recurso  $p$ . El estado inicial  $s_o$  representa el genotipo del individuo y está asociado a un fenotipo de reproducción  $\mathbf{F}_R$ . Cuando un individuo de genotipo  $G$  pasa por el estado  $\mathbf{F}_R$ , se agrega un nuevo individuo en la población. Estudiaremos cómo varía la dinámica poblacional con la similitud genética entre individuos y sus descendientes (bajo alguna métrica de similitud) y la herencia genética (en forma de energía) entregada en el nacimiento. Las tablas como la mostrada se generan al azar en cada realización y se mantienen constantes a lo largo de esta.

## 2.1. Un ejemplo simple

La evolución de los estados en este modelo, de acuerdo al ejemplo de la Tabla 2.1, se representa gráficamente en la Figura 2.1. Existe un único atractor para todos los estados iniciales, que es el ciclo  $010 \rightarrow 101 \rightarrow 110 \rightarrow 100$ . En general podrían existir más cuencas de atracción. Por ejemplo, si en el mapeo de la Tabla 2.1 cambiamos  $000 \rightarrow 100$  por  $000 \rightarrow 001$ , habría dos cuencas separadas en vez de una sola.

| $s_i$ | $s_{i+1}$ | $p(s_i)$ | $C(s_i)$ | $\mathbf{F}_R$ |
|-------|-----------|----------|----------|----------------|
| 000   | 100       | 0.8369   | 0.3628   | 000            |
| 001   | 000       | 0.2995   | 0.9326   | 001            |
| 010   | 101       | 0.3057   | 0.1626   | 000            |
| 011   | 101       | 0.2207   | 0.2925   | 100            |
| 100   | 010       | 0.6856   | 0.1192   | 110            |
| 101   | 110       | 0.8844   | 0.8525   | 100            |
| 110   | 100       | 0.6767   | 0.5595   | 100            |
| 111   | 010       | 0.1919   | 0.3335   | 011            |

**Tabla 2.1:** Ejemplo de mapeo de Kauffman para un código genético de 3 bits. Los individuos comienzan en un estado  $s_o = \mathbf{G}$ .



**Figura 2.1:** Árbol que representa la dinámica de los estados de la Tabla 2.1.

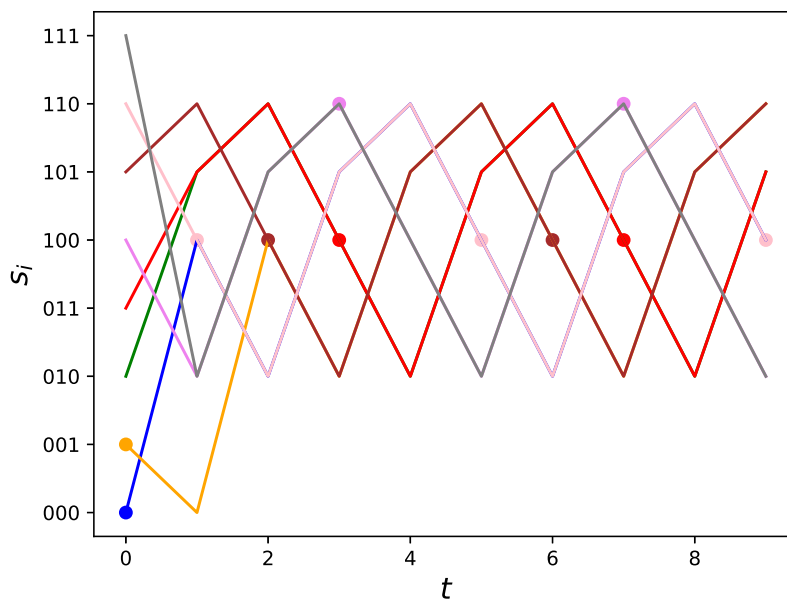
La Figura 2.2 representa la evolución de los 8 posibles genomas (estados iniciales) del sistema. Además, se señalan con puntos del mismo color los estados correspondientes a reproducciones, pero la dinámica de la descendencia no se muestra.

La Figura 2.3 muestra la evolución de la energía de los mismos individuos mostrados en la Figura 2.2. Estas evoluciones no tienen en cuenta que la energías deben ser positivas, sino que simplemente muestran los valores que toma la energía al pasar por los estados sucesivos de la Tabla 2.1 si el recurso fuera siempre suficiente.

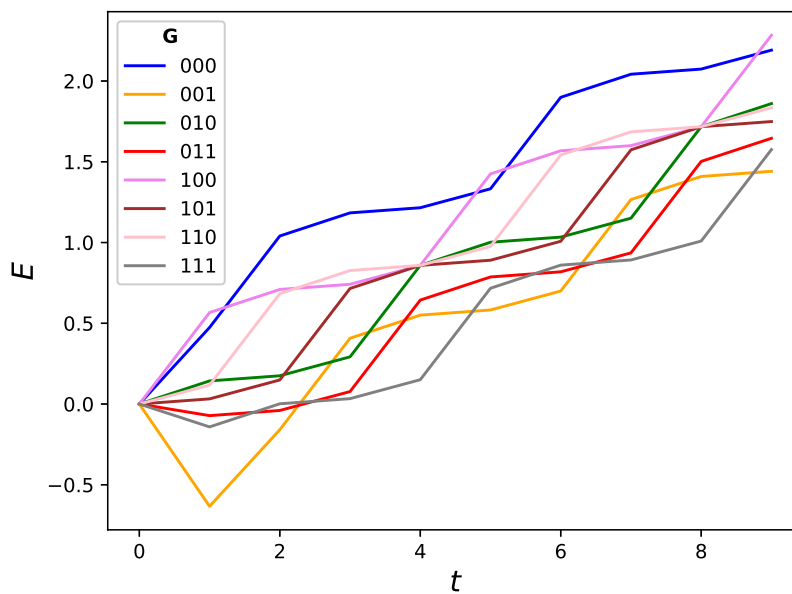
En la Figura 2.4 mostramos la evolución de una población que inicialmente tiene un individuo de cada genoma, según las reglas mencionadas para el caso particular del mapeo de la Tabla 2.1. Para simplificar la interpretación de los roles que cumplen los diferentes genomas, suponemos que los nuevos individuos tienen exactamente el mismo genotipo que sus padres (no hay errores de copia), y ninguno recibe energía. Se observan 4 comportamientos asintóticos diferentes, distinguidos por sus curvas de crecimiento. En primer lugar,  $\mathbf{G} = 100, 101,$  y  $110$  tienen comportamientos exponenciales (notar la escala logarítmica vertical). Los linajes con  $\mathbf{G} = 000$  y  $010$  tienen un comportamiento lineal (donde el último es constante e igual a 1) y el resto de los linajes se extinguen inmediatamente.

Los linajes que se extinguen inmediatamente son los que agotan su energía antes de reproducirse. Si el recurso es inagotable, esto se infiere de la Figura 2.3. En esta se ve que los linajes de  $\mathbf{G} = 011, 111,$  y  $001$  (amarillo, rojo y gris) tienen  $E < 0$  a  $t = 1$  y por lo tanto no sobrevivirán.

Los linajes  $\mathbf{G} = 000$  y  $010$  tienen comportamientos lineales ya que cumplen con una característica común. Ambos se reproducen fuera de ciclo atractor  $010 \rightarrow 101 \rightarrow 110 \rightarrow 100$ . En el caso de  $\mathbf{G} = 010$ ,



**Figura 2.2:** Estados sucesivos de los 8 individuos posibles de la Tabla 2.1. En puntos se señalan los estados donde ocurrirían reproducciones. Notar que algunas líneas que parecen desaparecer en el gráfico solo quedan ocultas cuando entran al ciclo atractor  $010 \rightarrow 101 \rightarrow 110 \rightarrow 100$ .

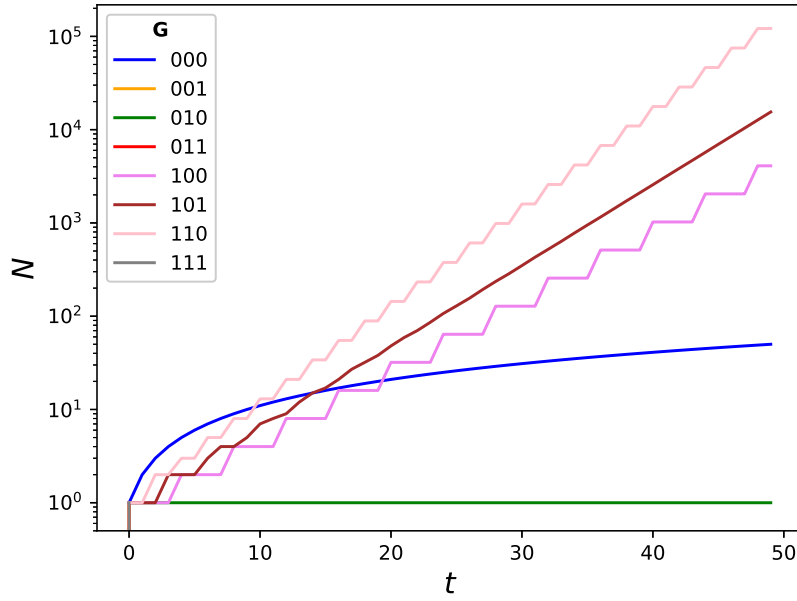


**Figura 2.3:** Evolución de la energía para los 8 individuos posibles de la Tabla 2.1. No se tiene en cuenta que si  $E < 0$ , los individuos mueren.

este se reproduce en el estado 000, pero como el mapeo determina que nunca pasa por este estado, la población se mantiene constante en  $N_{010} = 1$ . El linaje  $\mathbf{G} = 000$  se reproduce únicamente en el estado 000, es decir al nacimiento, y por lo tanto una vez en la vida. Si denotamos  $N_\tau(t)$  a la población del genotipo  $\mathbf{G} = 000$  de edad  $\tau$  a tiempo  $t$ , la dinámica de esta población puede resumirse en:

$$\begin{aligned} N_{\tau=0}(t) &= N_{\tau=0}(t-1), \\ N_{\tau>0}(t) &= N_{\tau>0}(t-1) + N_{\tau=0}(t-1). \end{aligned} \quad (2.1)$$

Denotemos  $N_{000}(t)$  a la población de la familia de genotipo  $\mathbf{G} = 000$  a tiempo  $t$ . La condición



**Figura 2.4:** Evolución de las poblaciones de cada uno de los genes para el modelo de redes de Kauffman de la Tabla 2.1.

inicial  $N_{000}(t) = N_{\tau=0}(0) = 1$  puede usarse con la primera línea de la Ecuación (2.1) para escribir  $N_{\tau=0}(t-1) = N_{\tau=0}(0) = 1$ . Esto implica que la población de esta familia crece linealmente:

$$\begin{aligned}
 N_{000}(t) &= N_{\tau=0}(t) + N_{\tau>0}(t), \\
 &= N_{\tau=0}(t-1) + N_{\tau>0}(t-1) + N_{\tau=0}(t-1), \\
 &= N_{000}(t-1) + N_{\tau=0}(0), \\
 &= N_{000}(t-1) + 1, \\
 \Rightarrow N_{000}(t) &= t + 1.
 \end{aligned} \tag{2.2}$$

Por otra parte, las dinámicas exponenciales de  $\mathbf{G} = 100, 101$  y  $110$  se entienden de la siguiente manera. La familia  $\mathbf{G} = 100$  se reproduce en  $\mathbf{F}_R = 110$ . En el presente ejemplo, la imagen de  $110$  también es  $100$ , por lo cual la reproducción transforma un individuo  $110$  en dos iguales de fenotipo  $100$ . Esto, a su vez, significa que un individuo  $100$  se transforma en dos del mismo fenotipo luego de 4 pasos de tiempo (el largo del ciclo). Podemos escribir, entonces,

$$\begin{aligned}
 N_{100}(t) &= 2N_{100}(t-4), \\
 N_{100}(t) &= 2^{t/4}.
 \end{aligned} \tag{2.3}$$

El linaje  $101$  presenta un período “transitorio” en el cual su crecimiento no puede escribirse de manera cerrada. Esto se expresa en las irregularidades del principio del crecimiento de la población de este linaje en la Figura 2.4. Sin embargo, como es de esperar por tener un fenotipo de reproducción dentro del ciclo atractor, su dinámica es exponencial con exponente entre el de la familia  $\mathbf{G} = 100$  y  $\mathbf{G} = 110$ .

Finalmente, el linaje  $\mathbf{G} = 110$  se reproduce en  $\mathbf{F}_R = 100$ . Cada dos pasos de tiempo, la población está compuesta únicamente por fenotipos  $010$  y  $110$  (y al tiempo siguiente lo estará solo por  $100$  y

101), por lo cual podemos reducir la dinámica fenotípica de esta familia a:

$$\begin{aligned} 110 &\xrightarrow{t+2} 010, \quad 110 \\ 010 &\xrightarrow{t+2} 110. \end{aligned} \quad (2.4)$$

Entonces, si la cantidad de individuos del linaje 110 con fenotipo  $s_i = 010$  a tiempo  $t$  es  $a_t$  y la cantidad con fenotipo  $s_i = 110$  es  $b_t$  (con  $N_{110}(t) = b_t + a_t$ ), estos evolucionan de acuerdo con:

$$\begin{aligned} b_t &= b_{t-2} + a_{t-2}, \\ a_t &= b_{t-2}, \\ \Rightarrow b_t + a_t &= b_{t-2} + a_{t-2} + b_{t-2}, \\ &= b_{t-2} + a_{t-2} + b_{t-4} + a_{t-4}. \end{aligned} \quad (2.5)$$

Así, la población evoluciona de acuerdo con la serie de Fibonacci, que cumple con la misma relación de recurrencia que la Ecuación (2.5).

## 2.2. Limitación de recurso

Hasta ahora consideramos que el recurso es inagotable para simplificar la dinámica poblacional. Sin embargo, el objetivo de este trabajo es estudiar el efecto de la selección natural sobre poblaciones, y una forma importante en la que esta actúa es usando el recurso para mediar interacciones.

Para implementar esto, diremos que el pedido de recurso  $p(s_i)$  de cada estado  $s_i$  determina la cota superior de la cantidad de recurso que consume cada individuo en ese estado, mientras que si este no alcanza (es decir,  $\sum_{i \in \mathcal{P}} p_i > f$ ), el recurso se reparte entre la población de acuerdo con pesos  $w_i$ . Así, la retribución  $\varphi_i = \min(p_i, w_i f)$  respeta

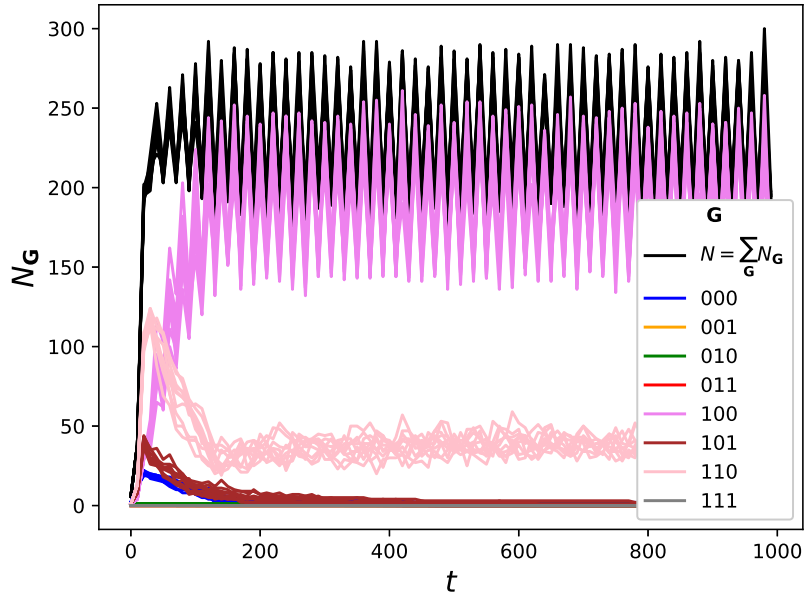
$$\sum_{i \in \mathcal{P}} \varphi_i \leq f. \quad (2.6)$$

En esta sección estudiaremos cómo cambian los resultados frente a diferentes dinámicas de recuperación  $df/dt = R(f)$ .

### Recuperación instantánea

En la Figura 2.5 mostramos cómo cambian las evoluciones de cada uno de los linajes mostrados en la Figura 2.4 cuando limitamos la cantidad de recurso disponible (usando nuevamente el mapeo de la Tabla 2.1). Para cualquier cantidad finita de recurso  $f(t)$ , la dinámica de recuperación es tal que  $f(t+1) = K$ . Es una recuperación instantánea hasta una capacidad de carga  $K$ . Para la Figura 2.5 hemos usado una capacidad de carga  $K = 100$ . Los pesos con los que se distribuye el recurso entre la población se eligen, a cada paso de tiempo, tomando  $N$  números al azar  $w_i$ . Luego, si  $\sum_i w_i > 1$ , se normalizan transformando  $w_i \rightarrow w_i / \sum_i w_i$ . Mostramos 10 realizaciones de la simulación para ilustrar los efectos del azar en este modelo. Este tipo de dinámica representa un sistema que recibe un flujo constante de energía, como los fotones que llegan del Sol y son procesados por organismos capaces de hacer fotosíntesis.

Lo primero que llama la atención de la Figura 2.5 es que el estacionario de la población total  $N$  (curva negra) llega mucho antes que el de las familias (curvas de colores). En el transitorio inicial ( $t < 200$ ) los linajes compiten entre sí por prevalecer, aumentando sus propias poblaciones a costa de las otras. El hecho de que en este forcejeo, la población total (restringida por el recurso y el consumo por individuo) se mantenga constante se deriva de que, al haber una sola cuenca de atracción, todos



**Figura 2.5:** Evolución de la población para el caso de la Tabla 2.1. Los pesos  $w_i$  se eligen al azar con distribución uniforme y se normalizan para que  $\sum_i w_i \leq 1$ . La recuperación del recurso es instantánea, asemejándose al flujo de fotones que llegan del Sol y los seres vivos con clorofila pueden procesar.

los linajes consumen la misma cantidad de recurso (en promedio). Por lo tanto, lo que determina la relación entre las poblaciones de los linajes no es la dinámica energética del grupo sino el consumo energético al nacimiento. Para la familia  $\mathbf{G} = 110$  es  $C(110) \approx 0.56$  mientras que para  $\mathbf{G} = 100$  es  $C(100) \approx 0.12$ , unas cinco veces menor. Esto coincide con la relación entre las poblaciones parciales y se entiende considerando que el consumo al nacimiento es el complemento de la probabilidad de sobrevivir (en el estacionario, cuando  $\sum_{\mathcal{P}} p \geq K$ ). Las otras dos familias con potencial para sobrevivir (según el escenario de recurso ilimitado) tienen consumos de aproximadamente 0.9, haciendo sus poblaciones un orden de magnitud por debajo de las otras. La restricción en la limitación del recurso hace que no se maximice la probabilidad de producir descendencia (notar que la tasa de reproducción de la familia  $\mathbf{G} = 110$  es mayor que la de 100), es decir que no se maximiza la *fitness* en su sentido convencional [39].

## Recuperación lineal

Consideremos ahora el caso en el que la dinámica de recuperación del recurso es lineal con tasa  $\alpha > 0$ ,

$$R(f) = \alpha f. \quad (2.7)$$

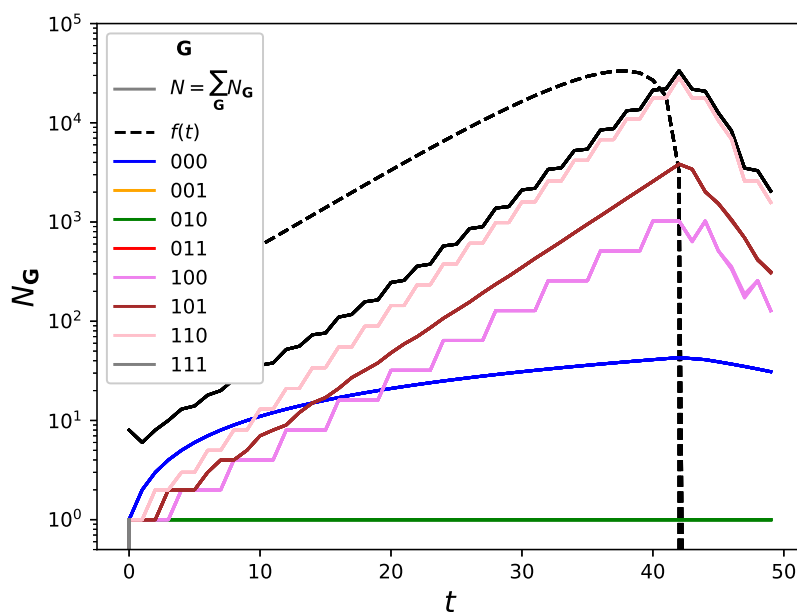
Supongamos que  $f(t) < \sum_{i \in \mathcal{P}} p_i(t)$ , donde  $p_i(t)$  es el pedido de energía de la criatura  $i$  a tiempo  $t$ . En este caso,  $f$  se distribuye entre la población, con lo cual  $f(t+1) = 0$  (sin importar la forma en la que se distribuye), que es un equilibrio (aunque inestable) de la Ecuación (2.7). En este caso, el recurso nunca se vuelve a recuperar y la población se extingue. Por otra parte, si el recurso es suficiente, la población no distingue entre este caso y el ilimitado. Tendrá una dinámica exponencial hasta que su ingesta sobrepase el recurso disponible (si la dinámica permite tal escenario) y luego se extinguirá.

Incluso si se introduce una cota inferior a la cantidad de recurso disponible para que este nunca llegue a cero, la población eventualmente se extinguirá porque el recurso no se puede recuperar lo suficientemente rápido para sostener a la población. Formalizaremos este problema, cuyo origen es la dinámica continua del recurso y la falta de un sistema de control de la reproducción, en la Sección 2.2.

En la Figura 2.6 se muestra la evolución de los mismos linajes que mostramos en las Figuras 2.4



y 2.5. Dependiendo de la tasa de crecimiento del recurso, hay dos regímenes posibles. Para tasas de recuperación bajas (con suficiente recurso inicial), la población presenta un crecimiento exponencial inicial y luego, cuando el consumo de la población supera el recurso disponible, agota el recurso y se extingue. Para tasas de recuperación mayores, el crecimiento tanto del recurso como de la población es exponencial en todo momento, y la población nunca se extingue. Esto es una condición irreal ya que, a pesar de que el recurso es finito para todo tiempo finito, crece indefinidamente, implicando que tiene alguna fuente de energía inagotable.



**Figura 2.6:** Evolución de la población para el caso de la Tabla 2.1 imponiendo una recuperación lineal del recurso en lugar de instantánea. Para tasas de recuperación bajas, como la que se muestra, el consumo de la población eventualmente supera la capacidad del recurso y se extingue. Para valores mayores de la tasa de recuperación, el sistema adopta una dinámica exponencial en recurso y población.

## Recuperación logística

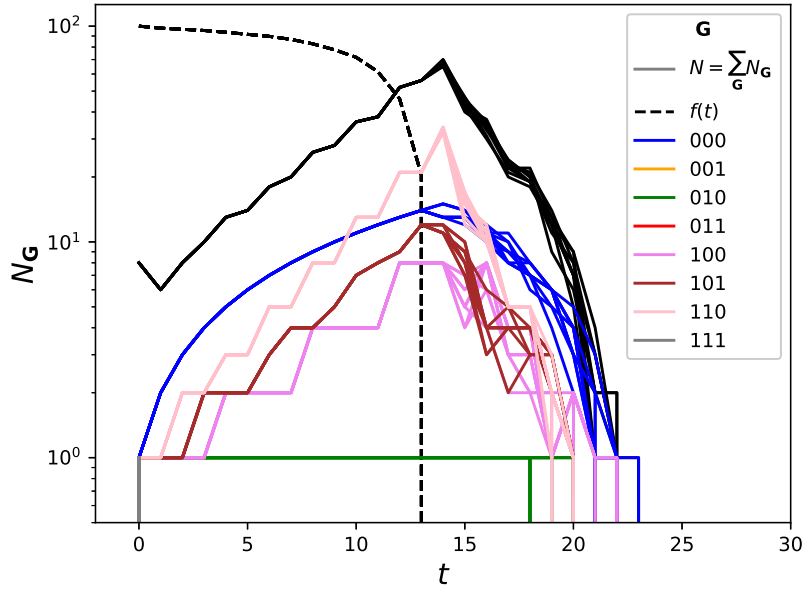
Finalmente, consideremos el caso en el cual el recurso se recupera logísticamente:

$$R(f) = rf \left(1 - \frac{f}{K}\right), \quad (2.8)$$

donde  $r$  es la tasa de recuperación y  $K$  la capacidad de carga. La Ecuación (2.8) define una función que crece exponencialmente para valores pequeños de  $f$  (despreciando  $f^2$ ) pero satura continua y asintóticamente al valor  $K$ . Destacamos que es una corrección de segundo orden respecto al crecimiento lineal de la Ecuación (2.7). Por esta razón, tanto matemática como conceptualmente, es la dinámica más simple que podemos usar si queremos mantener el realismo. En particular, no requiere energía infinita como el crecimiento lineal.

Por un argumento muy similar al que hicimos con el crecimiento lineal, podemos conjeturar que la población no sobrevivirá con esta dinámica. También podemos argumentar que si el consumo de la población supera al crecimiento exponencial (que no está acotado) para algunas tasas de recuperación, definitivamente saturará el crecimiento logístico (que sí está acotado).

Sin más, en la Figura 2.7 mostramos la dinámica de la población (tanto total como por linajes) y del recurso con este tipo de recuperación. Efectivamente, la población se extingue, esta vez sin importar la tasa de recuperación o la capacidad de carga.



**Figura 2.7:** Dinámica de la población (total y por familias) y el recurso cuando este se recupera logísticamente. La extinción que se observa no es producto de la elección de parámetros sino que ocurre para cualquier tasa de recuperación y capacidad de carga del recurso.

## Recurso con recuperación continua

A continuación mostraremos el problema que surge de un recurso con dinámica de recuperación continua. Más precisamente, supongamos que  $R(f)$  es continua en un entorno de  $f_o = 0$ , que también es un equilibrio de  $R$  (esto equivale a suponer que no hay generación espontánea del recurso [40, 41]). Si  $f_o$  es un equilibrio estable, el recurso se agotará dentro del entorno donde la dinámica de  $R$  es continua sin importar el efecto de la población. Por lo tanto, supondremos que  $f_o$  es un equilibrio inestable. Supongamos, además, que las tasas de reproducción individuales no dependen de la cantidad de recurso disponible. Esta última es la suposición menos realista de este resultado. A pesar de que seres vivos muy simples no obtienen información sensorial del recurso, podría argumentarse que obtienen información de manera indirecta. Incluso las moléculas prebióticas precursoras de la vida (individuos) no podían autocatalizarse (reproducirse) si no había nutrientes (recurso) [42]. Sin embargo, el resultado es de interés porque muestra la diferencia entre que sobrevivan los *más aptos* (si hay una sola dinámica, es la más apta) o los que *pueden*. Además, para seres vivos muy simples (por ejemplo, la reproducción podría ser  $\nu_{f_T}(f) = 0$  si  $f < f_T$  y  $\nu_{f_T}(f) = \nu$  en otros casos), el sistema de control puede ser demasiado lento (a nivel poblacional) para evitar las extinciones.

El problema surge de que, si los individuos no tienen reproducción controlada, no pueden distinguir entre recurso suficiente y recurso ilimitado, con lo cual el crecimiento poblacional será exponencial. Este crecimiento no es acotado y eventualmente superará la disponibilidad de recurso (que suponemos que se limita en algún momento para no implicar energía infinita). Como la recuperación  $R(f)$  es acotada, el consumo total  $N\langle p \rangle$  siempre superará a  $R(f)$  en un entorno del equilibrio inestable  $f_o = 0$  (asegurado por la continuidad) y hará que la dinámica total,

$$\frac{df}{dt} = R(f) - N\langle p \rangle, \quad (2.9)$$

tenga un equilibrio estable en  $f_o = 0$ . Notar que en el caso de recuperación inmediata, el hecho de que la dinámica sea discontinua invalida la hipótesis de que el consumo total supera a la recuperación cerca del equilibrio.

## 2.3. Errores genéticos

En la sección anterior consideramos individuos que generan copias perfectas de sí mismos al reproducirse. Como bien sabemos, durante la reproducción siempre ocurren errores de copia [43], con lo cual los descendientes nunca son idénticos a sus progenitores. Para implementar esto, introducimos una calidad de copia  $0 \leq c \leq 1$ . Esta representa la probabilidad de que un individuo nuevo en la población no mute su código genético respecto al del individuo que lo generó. De lo contrario, con probabilidad  $1 - c$ , la distribución de probabilidad es uniforme entre los  $2^L$  códigos genéticos distintos.

La decisión de que la distribución de la descendencia sea uniforme entre todos los genes y no, por ejemplo, decreciente con la distancia de Hamming entre las representaciones binarias de los genotipos, se basa en que los genes no tienen interpretación física. Sobre realizaciones del experimento, distintos genes podrán representar estados fenotípicos similares independientemente de este tipo de métricas. Tomamos, entonces, el modelo más simple que nos permita estudiar cómo repercute este cambio en las simulaciones.

A continuación, estudiaremos cómo evoluciona la población total y la de las distintas familias. Usaremos una recuperación instantánea de recurso debido a las extinciones halladas en las otras dinámicas.

### Calidad de copia constante

En primer lugar, estudiamos qué sucede cuando la calidad de copia es constante a lo largo de la simulación. Sabemos que, para  $c = 1$ , la evolución debería ser igual a la observada en la Figura 2.5. Por otro lado, para calidad de copia 0, la distribución de probabilidad de los genes de los descendientes es uniforme. Esto hace que cada individuo que se reproduce aporta uniformemente a las poblaciones familiares, por lo cual esperamos que las poblaciones de las familias sean aproximadamente iguales.

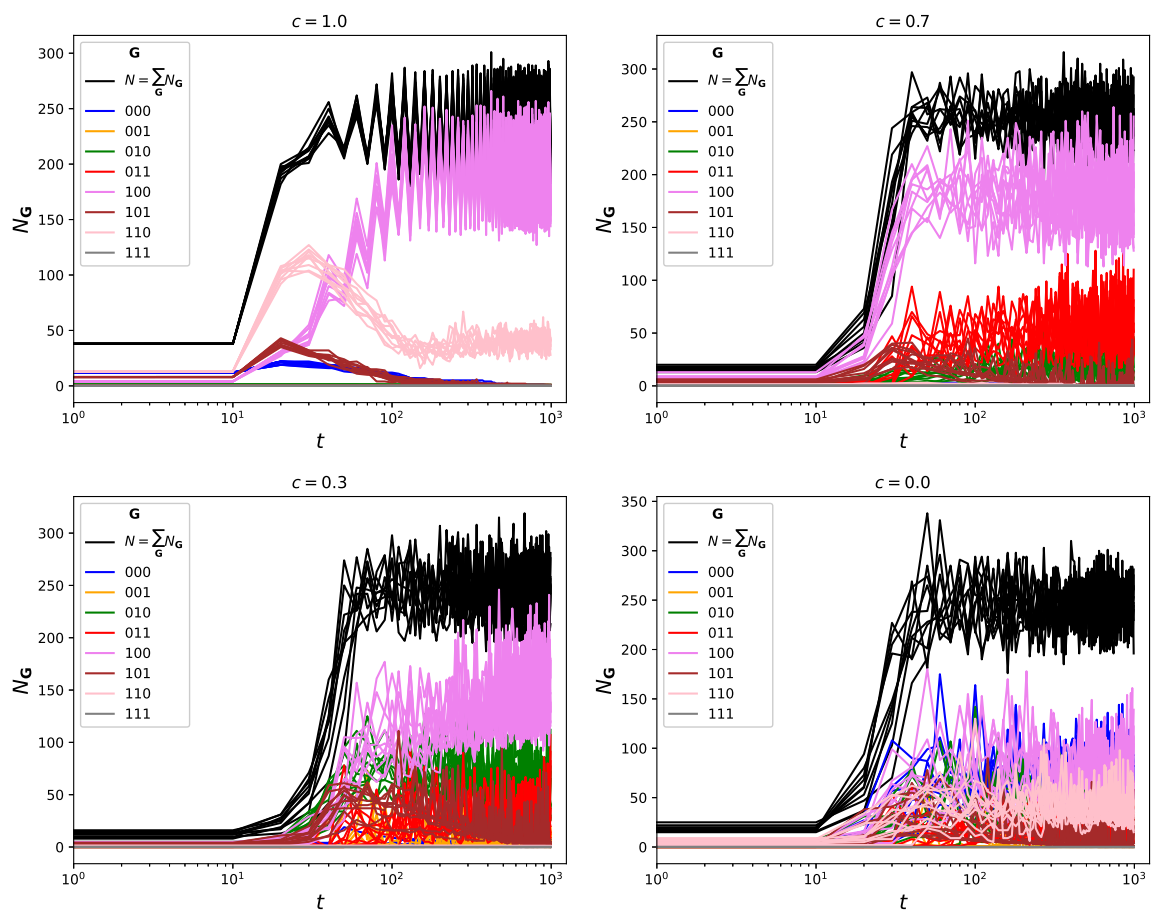
En la Figura 2.8 mostramos las evoluciones de las poblaciones para cuatro valores decrecientes de  $c$ . Efectivamente, observamos que en los casos límites, la evolución se comporta como esperamos por la argumentación anterior. Los casos intermedios sugieren que la transición de las poblaciones familiares para distintas calidades de copia es relativamente suave. Además, en todos los casos la población total es la misma, sugiriendo que esta está restringida únicamente por el recurso.

En la Figura 2.9 se muestran las poblaciones asintóticas de la Figura 2.8. Podemos observar que, en realizaciones, los resultados son bastante consistentes, indicando que los efectos estocásticos no son muy grandes. Por otro lado, los puntos parecen tener comportamientos continuos (sin saltos grandes), aunque, en general, no monótonamente crecientes o decrecientes. Para calidad de copia nula, las poblaciones no son exactamente iguales como esperábamos pero sí presentan mínima dispersión. La población total, por otra parte se mantiene relativamente constante.

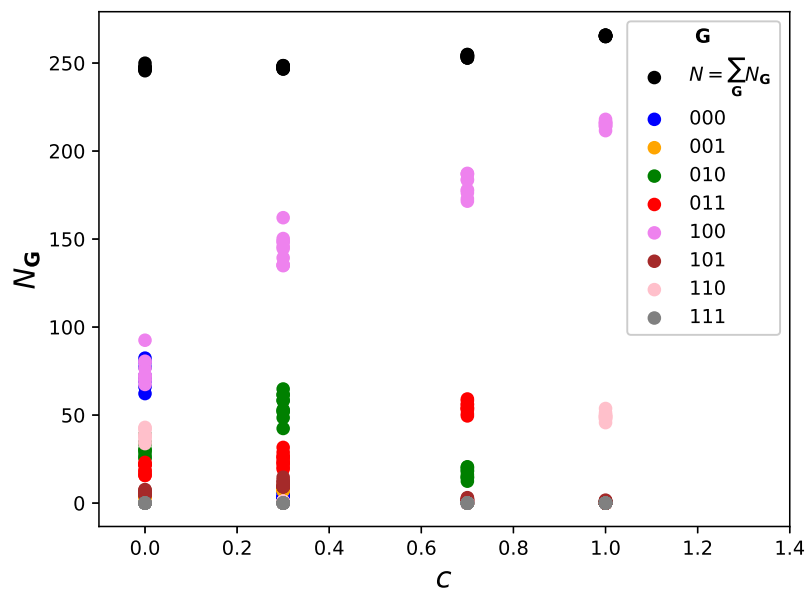
### Calidad de copia libre

Los resultados de la Figura 2.8 son interesantes desde un punto de vista didáctico. Sin embargo, sería más realista que la calidad de copia pueda evolucionar, ya que los mecanismos de corrección de errores del código genético están contenidos en el mismo código genético y por lo tanto sujetos a la selección natural.

Para implementar esto, introducimos una calidad de copia por individuo  $c_i$ . Esta calidad de copia comienza con un valor pequeño,  $c \approx 0$ , con un parámetro de fluctuación de la calidad de copia universal  $\Delta_c$ . Generación tras generación, las mutaciones ocurren según  $c_i$ , pero moduladas según el parámetro de la fluctuación. Si el descendiente de  $c_i$  es  $c_j$ ,  $c_j$  se elige de una distribución uniforme  $|c_j - c_i| \leq (1 - c_i)\Delta_c$ . El factor  $(1 - c_i)$  hace que la calidad de copia también contemple que la calidad



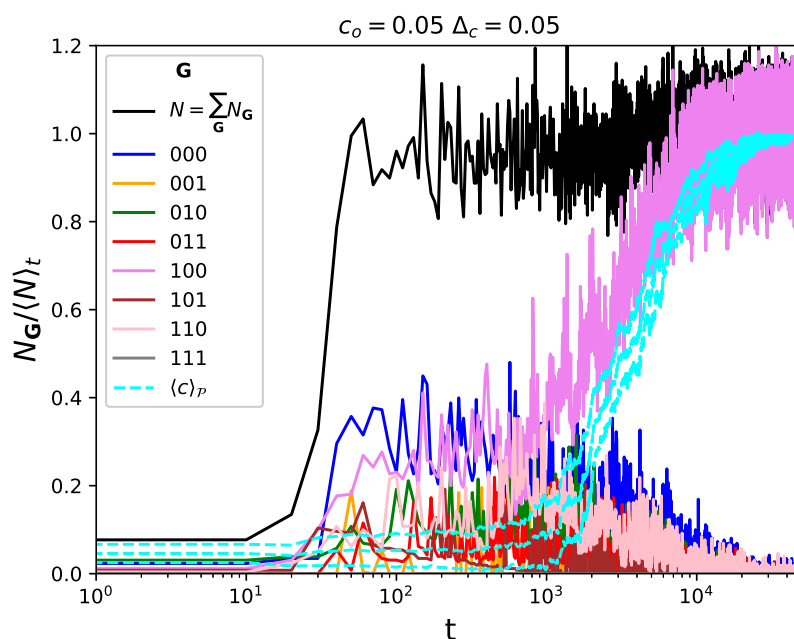
**Figura 2.8:** Evolución de la población total y de cada una de las familias genéticas para distintas calidades de copia  $c = 1, 0.7, 0.3$  y  $0$ . Se observa que el caso límite de calidad de copia  $c = 1$  coincide con los resultados de la Figura 2.5 mientras que el límite  $c = 0$  produce una distribución aproximadamente uniforme entre las familias. Por otro lado, la población total es la misma para cualquier calidad de copia, sugiriendo que ésta depende solo de la cantidad de recurso disponible.



**Figura 2.9:** Poblaciones asintóticas obtenidas de la Figura 2.8. Mientras que la población total se mantiene constante, las poblaciones parciales varían continua pero no monótonamente.

de copia sea similar en los descendientes cuando esta sea alta en los padres, mientras que  $\Delta_c$  modula el tamaño de estas fluctuaciones.

En la Figura 2.10 se muestra la evolución de la población y su composición (normalizada a la población total en el estacionario), así como la evolución de la calidad de copia media de la población. Se utilizó un parámetro de fluctuación  $\Delta_c = 0.05$  y una calidad de copia inicial  $c_o = 0.05$ . Se observa que las poblaciones comienzan de manera muy similar a cuando la calidad de copia es baja y constante (uniformemente distribuidas). Sin embargo, en algún momento ocurre una ruptura de simetría y una de las familias comienza a prevalecer por sobre las otras (curva magenta). Dicha familia es la misma que tiene la mayor población en los experimentos con calidad de copia constante. Una vez que la diferencia se marca un poco, el sistema transiciona a un estado en el que solo una familia sobrevive, incluso sin la coexistencia que se observaba para calidad de copia constante e igual a uno.



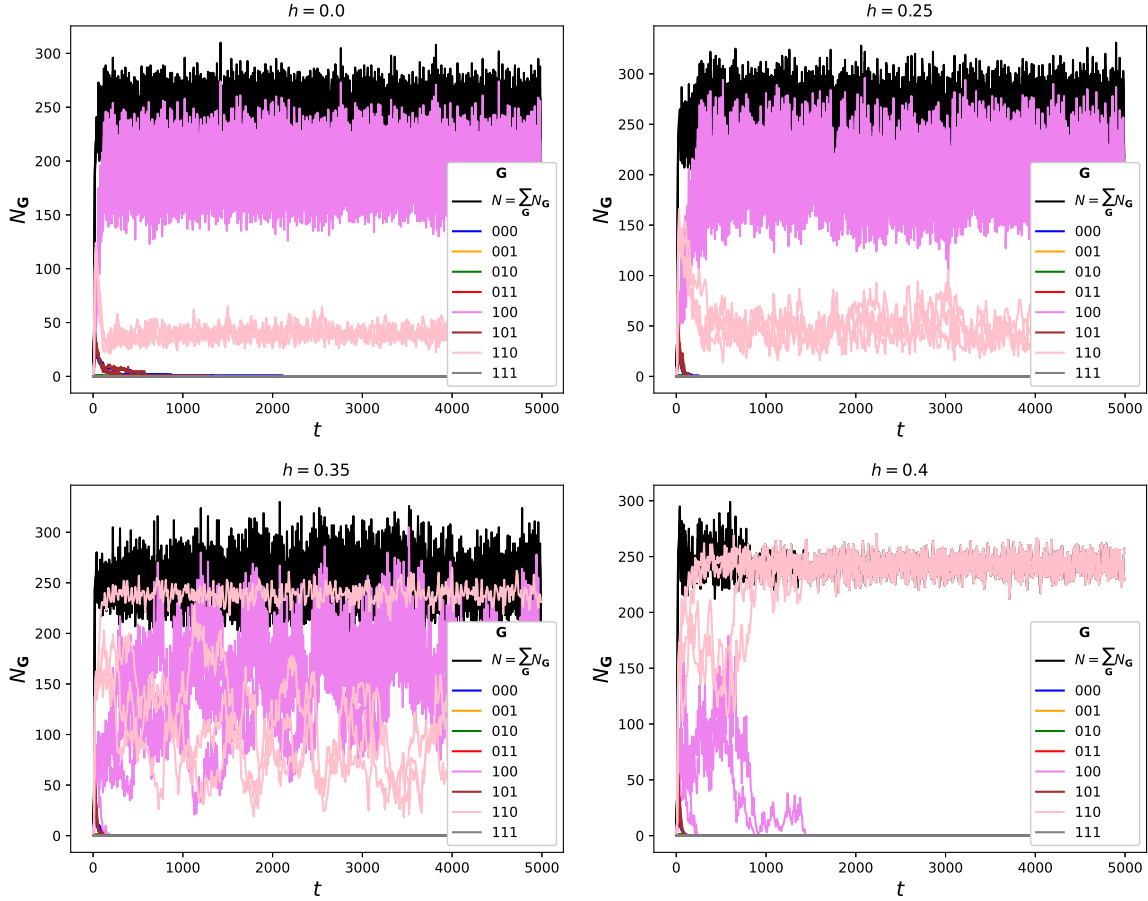
**Figura 2.10:** Evolución de la población normalizada a la población en el estacionario y poblaciones parciales. También se muestra la calidad de copia media de la población con su dispersión. Cuando empiezan a prevalecer los individuos que componen la mayor parte de la población en los experimentos de calidad de copia constante, la calidad de copia aumenta rápidamente, convergiendo a 1, y el resto de las familias se extinguen. Es interesante que el estacionario de la población total llega mucho antes de que suceda esto y no se ve afectada por los cambios en su composición por familias.

## 2.4. Herencia

Hasta ahora consideramos que no hay costo de reproducción y que los individuos nuevos no tienen ninguna diferencia o ventaja respecto a las primeras generaciones. No hay una “herencia” en el sentido de que no se transmite información salvo la genética, que generación tras generación se perfecciona a través de la selección natural. A continuación, liberaremos esta suposición y veremos cómo influye en la dinámica poblacional. Implementamos esto permitiendo que los individuos recién nacidos reciban de sus padres una fracción  $h$  (que depende del progenitor) de su energía, dando una ventaja al momento nacimiento. Como la energía debe conservarse, esto supone una desventaja para los progenitores. Casos análogos a este tipo de sacrificio altruista en la realidad podrían incluir enseñar a cazar, compartir alimentos, o incluso el costo del período de gestación.

En la Figura 2.11 mostramos cuatro realizaciones de simulaciones con calidad de copia  $c = 1$ . Nos

restringimos a este valor porque, como observamos en la Figura 2.10, la calidad de copia libre tiende a esto. Cada panel corresponde a una fracción de energía  $h$  que recibe el recién nacido, como se indica sobre el recuadro.



**Figura 2.11:** Evolución de la población total y las familias que lo componen para distintas fracciones de energía heredada  $h$

En la Figura 2.11 podemos observar que, a medida que aumenta la fracción de energía heredada, el tamaño de las fluctuaciones en las poblaciones de cada familia crece. Para  $h = 0.4$  se observa incluso que la familia 100, hasta entonces siempre predominante en población, se extingue. Esto da lugar a que la familia 110, que tomaba el segundo lugar, llene la población. Esto indica que los miembros de la familia 100, a pesar de ser los más aptos en las situaciones anteriores, son mucho más sensibles a fluctuaciones de recurso. Efectivamente, cuando un miembro de la familia 100 se reproduce, lo hace con un fenotipo  $s_i = 110$ . Este fenotipo produce (en el mejor de los casos) un cambio energético  $p(110) - C(110) = 0.12$ . La familia 110, por otro lado, se reproduce en 100, que tiene un cambio energético de 0.57, lo cual lo hace mucho más resistente a los costos de la reproducción.

Hasta ahora estudiamos siempre el ejemplo de la Tabla 2.1. Con este único ejemplo de tres bits, y haciendo pequeñas modificaciones al modelo, hallamos numerosos resultados, tanto esperados como no. Al aumentar el largo de la cadena de ADN, esperábamos encontrar comportamientos nuevos, pero en casi todos los aspectos reflejaron la fenomenología que acabamos de discutir. Una diferencia que notamos es que la coexistencia de especies en el estado estacionario con recurso limitado, como se observa en la Figura 2.5, no sucede a menudo. Incluso se vuelve más raro a medida que se aumenta el largo de la cadena, presumiblemente por la probabilidad de que se formen más atractores en el mapeo. Sin embargo, reconocemos que la capacidad computacional disponible (limitada por el crecimiento

---

exponencial en los tamaños de los mapeos) permitió simular hasta  $L = 15$ , mientras que las células más simples suelen tener miles de genes. Debido al carácter poco estocástico de este modelo, a continuación propondremos otro, que utiliza valores medios de las expresiones fenotípicas para centrarnos en los efectos de la selección natural.





## Capítulo 3

# Modelo de información

*“But they don’t know that we know they know we know.”*

— Phoebe Buffay

El modelo de redes de Kauffman presentado en el Capítulo 2 es simple, sin mucha interpretación, y su único objetivo es presentar los modelos de selección natural basados en energía. En particular, no da mucha información acerca del rol de la selección natural en sí. Una de sus mayores desventajas es que pequeños cambios en el genotipo producen grandes cambios en el fenotipo. En la realidad esto puede ocurrir (una persona con síndrome de down tiene un cromosoma menos o más, que contiene muchísimos genes, pero es indudablemente un ser humano) pero generalmente no es así, sino que individuos de genotipo similar resultan similares (parientes cercanos son físicamente similares). Esto permite que, a través de pequeñas mutaciones, las poblaciones puedan explorar nuevas posibilidades para resolver sus necesidades vitales pero “aprendiendo” (estrictamente, incorporando información mediante la selección natural) de los errores de sus antepasados. Con el objetivo de estudiar esta forma de evolución, presentamos en este capítulo una versión simplificada de los modelos de Kauffman que provee alguna intuición sobre el proceso.

### 3.1. Transmisión vertical de información

La principal modificación que introducimos en el modelo de Kauffman se basa en que, en los ciclos a los que convergen los estados de un mapeo de Kauffman, los valores fenotípicos promedio (sobre un período) están bien definidos y son constantes. Esto sugiere reemplazar el fenotipo  $\mathbf{F} = s_i$  y su pedido de recurso  $p$  y consumo energético  $C$  asociados, por un fenotipo constante  $\mathbf{F} = (p, C, \nu, h)$ , donde  $\nu$  y  $h$  son la tasa de reproducción y la fracción de energía heredada de un individuo respectivamente, y cumplen con  $0 \leq \nu, h \leq 1$ . Debido a las extinciones halladas para dinámicas continuas de recurso en el Capítulo 2, supondremos una recuperación instantánea del recurso  $f(t+1) = K$  para todo  $f(t)$ . A cada instante de tiempo se asigna a los individuos, al azar y uniformemente, un orden de ingesta, que representa la prioridad sobre la adquisición del recurso. En caso de que el recurso disponible para ese individuo sea mayor que el pedido de energía, solo se ingiere la cantidad pedida. De lo contrario, ingiere lo disponible. Así, podemos escribir la dinámica energética como:

$$\frac{dE}{dt} = \min(p, wK) - C - Eh\delta_\nu(t), \quad (3.1)$$

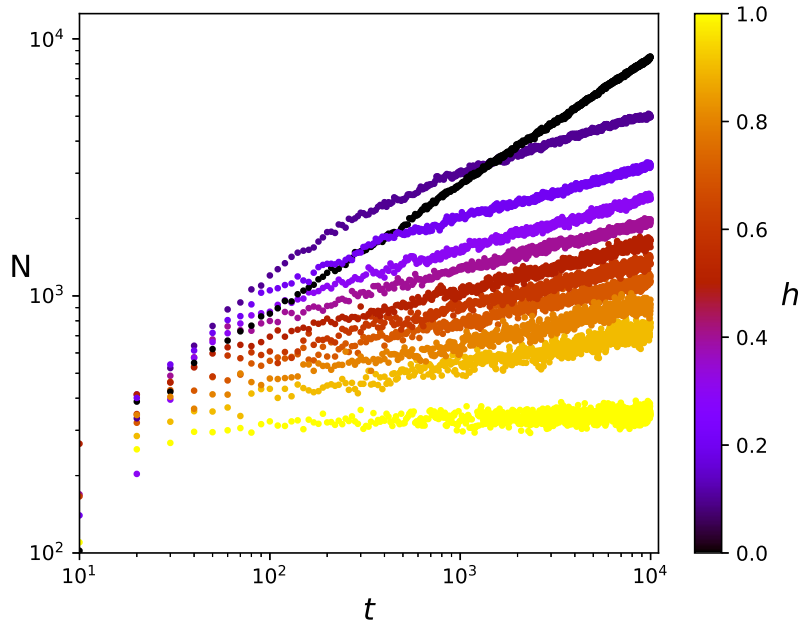
donde  $\delta_\nu(t)$  es un proceso de Poisson que vale 1 con tasa  $\nu$  y 0 en otros casos, y  $w$  es la fracción de recurso remanente después de la ingesta de los individuos anteriores. Notar que el hecho de que la

retribución sea  $\varphi = \min(p, wK)$  y no  $\varphi = \min(p, wf)$  es consecuencia de la dinámica del recurso que usamos.

Por simplicidad, la información genética que codifica el fenotipo en este modelo está contenida en un genotipo compuesto por los mismos elementos que el fenotipo. Además, contiene el parámetro de mutación  $c$  que indica la calidad de copia, con lo cual  $\mathbf{G} = (p, C, \nu, h, c)$ . Las mutaciones entre una generación y la siguiente ocurren con probabilidad uniforme centrada en los genes de los padres y ancho  $1 - c$  para cada uno de los componentes (restringido al dominio  $[0, 1]$ ). Salvo porque todos dependen de la misma calidad de copia, los genes varían independientemente. Esta suposición es equivalente a que todos los genes tienen el mismo mecanismo de copia y corrección y que los genes codifican las características fenotípicas de manera independiente. Ninguna de estas dos suposiciones es cierta en general [44], pero el caso real se reducirá al independiente y descorrelacionado mediante la diagonalización (o descomposición en bloques) de la matriz de correlación de los genes (tema que no trataremos en este trabajo).

### Sin transmisión de información

Comenzamos estudiando la evolución del sistema cuando la calidad de copia de los individuos es constante y nula a lo largo de la simulación. La Figura 3.1 muestra la evolución de la población total para distintas fracciones de energía heredadas  $h$ , constantes a lo largo de cada evolución y con el resto del genotipo distribuido uniformemente en su dominio.



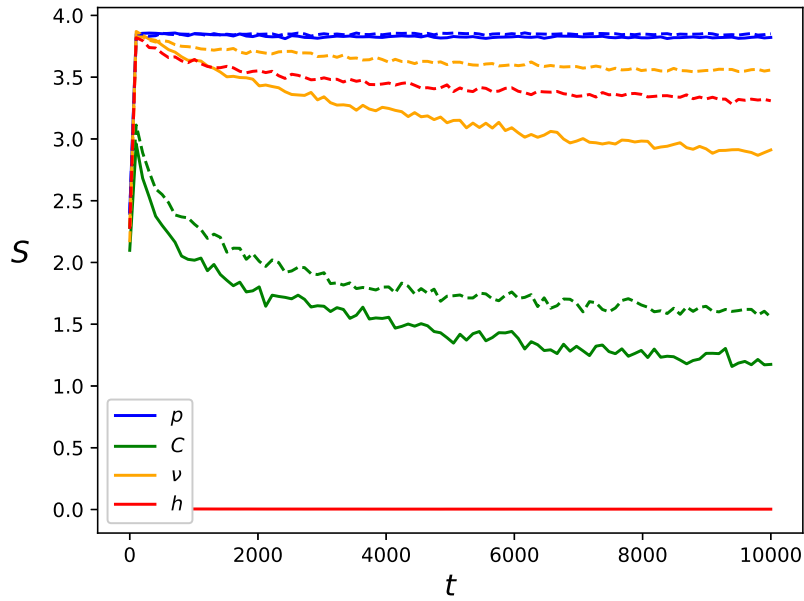
**Figura 3.1:** Evolución temporal de la población para distintas fracciones de energía heredadas  $h$ , constantes en cada simulación. Se observa que la población evoluciona algebraicamente en el tiempo  $N \propto t^\gamma$  a partir de  $t \approx 1000$  con distintos exponentes  $\gamma$  según el valor de la fracción de energía heredada.

En esta figura observamos que las poblaciones evolucionan, a tiempos largos, de acuerdo a una relación algebraica  $N \propto t^\gamma$  (nótese la escala doble logarítmica). A medida que aumenta la fracción de energía heredada, disminuye el exponente  $\gamma$  de la población correspondiente. Además, en todo el rango intermedio  $0.1 < h < 0.9$ , el exponente es aproximadamente constante mientras que en los extremos cambia abruptamente pero de manera continua, como mostraremos más abajo (Figura 3.5).

En el sentido matemático de la información,  $S = -\sum p \log(p)$ , y a la que nos referiremos como entropía para facilitar su distinción, podemos hacer la siguiente interpretación. Cuando no hay

transmisión de información de generación en generación, la entropía de la distribución genética en el nacimiento es máxima (la distribución es uniforme en el dominio). En cada generación el sistema disminuye la entropía de esta distribución mediante la selección natural (los individuos que han sobrevivido presionan a los recién nacidos para que sean al menos igual de aptos que ellos). Esto aumenta el orden en el sistema *quitando* información (disminuyendo la entropía). Esto es a costa del flujo de energía que se recibe como recurso y la muerte de muchos individuos que se encuentran fuera de los parámetros genéticos capaces de sobrevivir. Esto puede entenderse como que el sistema busca disminuir la energía “libre” (en forma de recurso) a costa de disminuir la entropía. La energía se almacena en los cuerpos de los individuos tomando una forma menos útil (solo puede ser aprovechada si se transmite a las generaciones siguientes), y representa un “objetivo” de la vida de degradar la energía disponible [35].

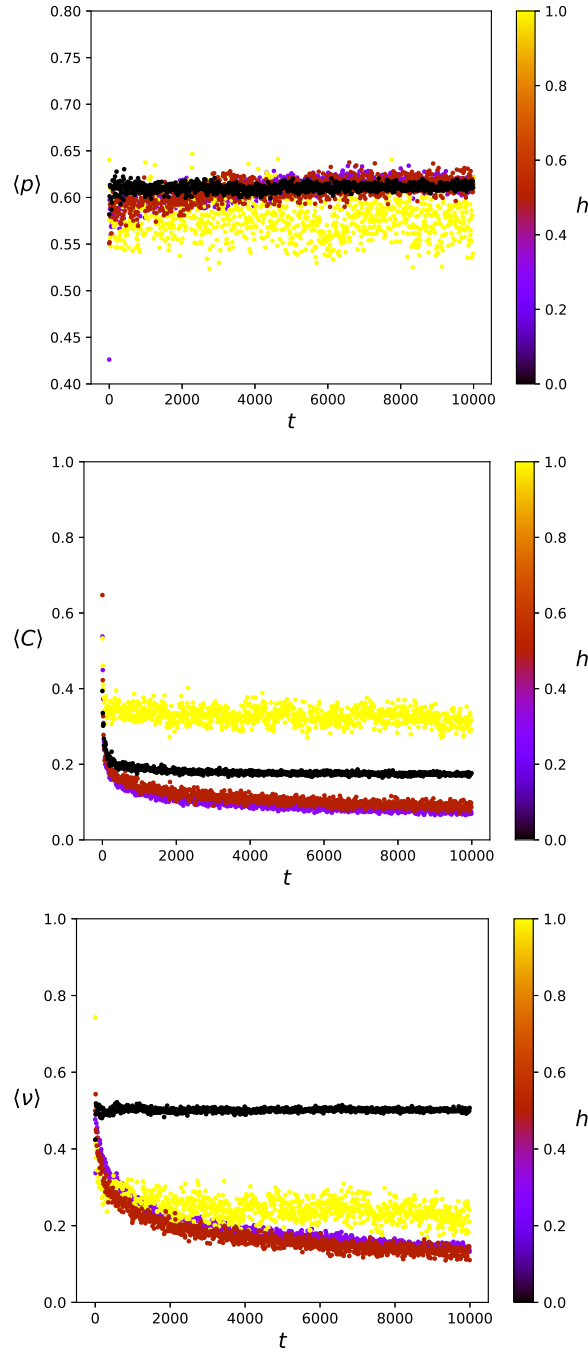
En la Figura 3.2 se muestra la evolución de las entropías correspondientes a las distribuciones de cada uno de los genes. En líneas llenas se muestran las de fracción de energía heredada constante ( $h = 0.2$ ) y en líneas de trazos las de fracción de energía heredada variable. Se observa que las segundas son siempre mayores que las primeras. La entropía total (igual a la suma ya que son variables independientes) también disminuye y también es menor para el caso constante que el variable.



**Figura 3.2:** Evolución temporal de las entropías de las distribuciones de cada uno de los genes para el caso de fracción de energía heredada constante ( $h = 0.2$ ) y variable. Para todos los genes, el caso variable tiene mayor entropía que el constante, pero la entropía total de ambos (igual a la suma por ser variables independientes) es decreciente. Esto es a costa del consumo del recurso, que actúa como una fuente de energía externa y permite aumentar el orden del sistema.

En la Figura 3.3 mostramos la evolución de los valores medios del pedido de recurso  $p$ , el consumo energético  $C$  y la frecuencia de reproducción  $\nu$  para las mismas simulaciones que se mostraron en la Figura 3.1. A pesar de que estas figuras no se muestran en escala doble logarítmica, vale la pena destacar que todas las curvas tienen un comportamiento algebraico como el de la población. El objetivo de mostrar las curvas en escala lineal es poder observar los valores asintóticos que se alcanzan como resultado de la presión evolutiva.

En estos gráficos, lo primero que notamos es que el pedido de recurso es prácticamente constante a lo largo de la simulación. Esto implica que la presión evolutiva sobre este gen es prácticamente nula comparada con la de los otros genes. En particular, para los valores extremos de la fracción de energía heredada,  $h = 0$  y  $h = 1$ , esta presión es incluso menor que la de valores de  $h \in (0, 1)$ , donde el valor



**Figura 3.3:** Evolución de los valores medios de el pedido de recurso  $p$ , el consumo energético  $C$  y la frecuencia de reproducción  $\nu$  para las mismas simulaciones que se mostraron en la Figura 3.1

medio de este gen crece levemente a lo largo de la evolución del sistema. El hecho de que el pedido de recurso sea un gen neutro (es decir que no haya presión genética sobre el mismo) puede parecer extraño en un principio, pero en realidad el valor restrictivo es el del consumo energético  $C$ . Siempre que la retribución  $\varphi = \min(p, wK)$  promedio compense los costos energéticos y de reproducción, el individuo sobrevivirá y producirá descendencia. A continuación formalizaremos esta idea.

Para una población con  $N$  individuos de pedido de recurso medio  $\langle p \rangle$ , la probabilidad de recibir la cantidad de recurso pedida  $p$  es  $M/N = K/N\langle p \rangle$ , donde  $M$  es el número de individuos que quedan en posiciones (en el orden de ingesta) tales que el recurso todavía no se agotó. Despreciaremos el (único) individuo que consume la cantidad remanente en vez de la pedida y supondremos que, en cualquier

otro caso, la retribución es nula. Así, la energía de un individuo que produce descendencia a un tiempo  $t$  es

$$E(t) = E_o + \left( \frac{K}{N\langle p \rangle} - \frac{C}{p} \right) pt = E_o + \frac{\nu}{\lambda} t, \quad (3.2)$$

(con una definición adecuada de  $\lambda$ ) y con probabilidad

$$\begin{aligned} \rho(E|E_o, \mathbf{G}, \mathcal{P}) &= \rho(t|E_o, \mathbf{G}, \mathcal{P}) \frac{dt}{dE}, \\ &= \exp \left[ -(E - E_o) \frac{\nu}{p \left( \frac{K}{N\langle p \rangle} - \frac{C}{p} \right)} \right] \frac{\nu}{p \left( \frac{K}{N\langle p \rangle} - \frac{C}{p} \right)}, \\ &= \lambda \exp(-(E - E_o)\lambda). \end{aligned} \quad (3.3)$$

Así, la distribución de probabilidad de la energía de nacimiento  $E'_o = Eh$  de un descendiente está dada por

$$\rho(E'_o) = \rho(E|E_o, \mathbf{G}, \mathcal{P}) \frac{1}{h}, \quad (3.4)$$

que finalmente permite determinar la probabilidad de producir descendencia fértil, comúnmente aceptada como *fitness* en un sentido biológico estricto, y que es el parámetro que proponen optimizar la mayoría de los modelos de selección natural [39]. Dejaremos expresada la forma integral de la solución para mostrar que puede obtenerse analíticamente, pero en modelos más complicados, la expresión será mucho más difícil de escribir y el método probablemente no sea útil.

$$\begin{aligned} \text{Prob} &= \int_{E'>0} \int_{\mathbf{G}'} \int_{E'_o>0} \rho(E'|E'_o, \mathbf{G}', \mathcal{P}') \rho(\mathbf{G}', \mathcal{P}'|G, \mathcal{P}) \rho(E'_o) dE'_o d\mathbf{G}' dE', \\ &= \int_{E'>0} \int_{\mathbf{G}'} \int_{E'_o>0} \exp(-(E' - E'_o)\lambda') \exp\left(-\left(\frac{E'_o}{h} - E_o\right)\lambda\right) \frac{\lambda\lambda'}{h} dE'_o d\mathbf{G}' dE', \\ &= \int_{E'>0} \int_{\mathbf{G}'} \int_{E'_o>0} \exp(-(E'\lambda' - E_o\lambda)) \exp\left(-E'_o\left(\frac{\lambda}{h} - \lambda'\right)\right) \frac{\lambda\lambda'}{h} dE'_o d\mathbf{G}' dE'. \end{aligned} \quad (3.5)$$

Debemos considerar que para producir descendencia fértil deben ocurrir dos cosas sucesivas:

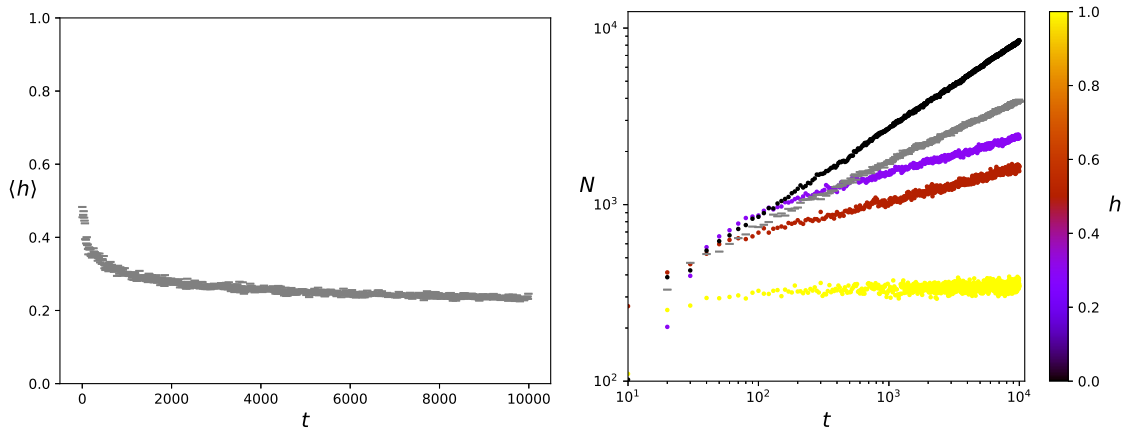
1. Un individuo debe reproducirse antes de que se le agote la energía. En caso de  $\lambda > 0$  ( $dE/dt > 0$ ), esto implica  $E'_o \in [0, \infty)$ . En caso de  $\lambda < 0$ ,  $E'_o \in [0, E_o]$ .
2. El nuevo individuo debe reproducirse antes de que se le agote la energía. Esto requiere de lo mismo que para el primero, pero ahora debe marginarse sobre los genes del nuevo individuo. Si  $\lambda' > 0$ ,  $E' \in [E'_o, \infty)$  y si  $\lambda' < 0$ ,  $E' \in [0, E'_o]$ .

Estas condiciones permiten hallar los límites de integración sobre las distintas variables que se integran en la Ecuación (3.5). Destacamos que, para que  $\lambda > 0$ , lo que se requiere es que  $C < Kp/N\langle p \rangle$ . Como hemos observado que  $N$  crece algebraicamente, y en  $t \approx 100$  ya es del orden de  $N \approx 1000$ , supondremos que el valor donde  $\lambda$  cambia de signo es en  $C = 2K/N$  (tomando  $p = 1$  y  $\langle p \rangle = 0.5$ ). Así, para  $\lambda > 0$ , obtenemos:

$$\begin{aligned} \text{Prob} &= \int_{E_o}^{\infty} \int_{p', \nu' \in [0, 1]^2} \left\{ \int_{E'_o}^{\infty} \int_{C' \in [0, \frac{2K}{N}]} \exp(-(E'\lambda' - E_o\lambda)) \exp\left(-E'_o\left(\frac{\lambda}{h} - \lambda'\right)\right) \frac{\lambda\lambda'}{h} dC' dE' \right. \\ &\quad \left. + \int_{E'_o}^0 \int_{C' \in [\frac{2K}{N}, 1]} \exp(-(E'\lambda' - E_o\lambda)) \exp\left(-E'_o\left(\frac{\lambda}{h} - \lambda'\right)\right) \frac{\lambda\lambda'}{h} dC' dE' \right\} dp' d\nu' dE'_o, \end{aligned} \quad (3.6)$$

y para  $\lambda < 0$  el límite superior de la última integral  $E'_o = \infty$  se debe reemplazar por  $E'_o = 0$ .

A continuación, mostramos lo que sucede cuando la fracción de energía heredada no es constante, sino que muta de una generación a la otra como los otros genes. En la Figura 3.4 mostramos la evolución del valor medio de este gen y la evolución de la población total (curvas grises), comparando esta última con algunas de las evoluciones con fracción de energía heredada constante. En el panel de la izquierda se observa que  $\langle h \rangle$  decae muy lentamente. En el régimen observado, la evolución del sistema se mantiene próxima a un valor  $\langle h \rangle \approx 0.2$ . A pesar de esto, la dinámica poblacional del sistema (panel derecho) no tiende al comportamiento de una simulación con el mismo valor constante (no es paralela a la curva violeta ni a la negra). Esto indica que, aunque los individuos no tienen memoria genética ni individual, el sistema genera un comportamiento colectivo que sí parece tenerla. Lo mismo ocurre con el comportamiento de los valores medios de la Figura 3.3, donde la evolución con una fracción de energía heredada variable se comporta de manera distinta a cuando esta es constante.

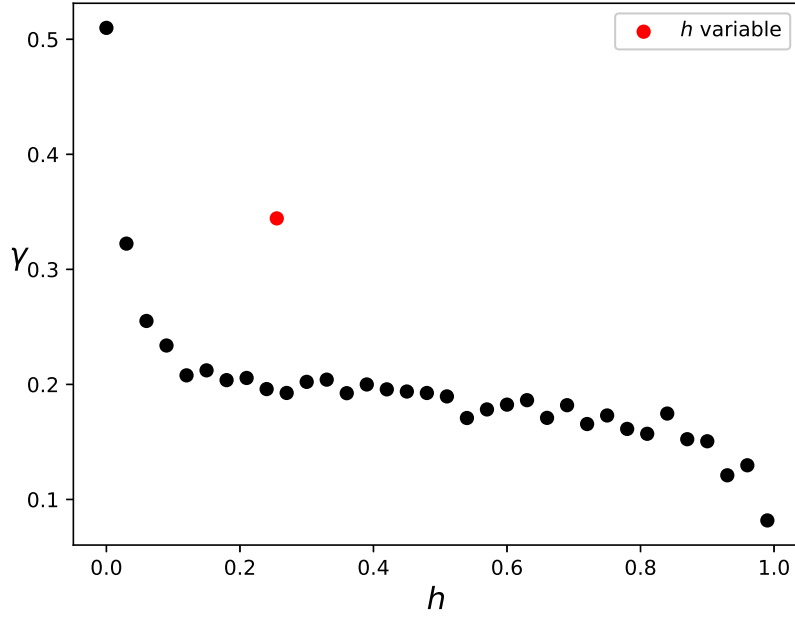


**Figura 3.4:** Evolución de la fracción de energía heredada promediada sobre la población y evolución de la población con fracción de energía heredada  $h$  variable, comparada con algunas donde la fracción de energía heredada es constante. Podemos observar que el valor medio de la fracción de energía heredada tiende a  $\langle h \rangle \approx 0$ , pero que en la escala de tiempo mostrada pareciera mantenerse en un valor  $\langle h \rangle \approx 0.2$ . Sin embargo, el comportamiento asintótico de  $N(t)$  tiene una pendiente distinta de la curva correspondiente para un valor de  $h$  realmente constante (curvas violeta y negra).

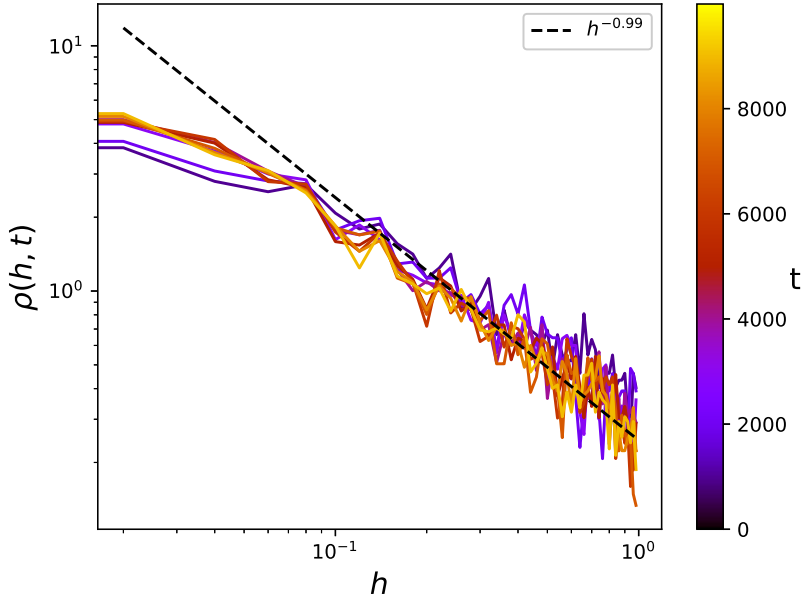
En la Figura 3.5 mostramos los exponentes  $\gamma$  ajustados de la evolución de la población para  $t > 1000$ , régimen lineal en escala doble logarítmica. Por un lado, disminuyen con la fracción de energía heredada cuando esta última es constante. Los exponentes presentan una región aproximadamente constante para  $0.2 < h < 0.8$ , mientras que en los extremos cambia abrupta pero continuamente, como anticipamos en la discusión de la Figura 3.3. Por otro lado, cuando la fracción de energía heredada es variable, el exponente efectivo de la población es mucho mayor (punto rojo) que el que le correspondería a su valor medio si este fuese constante.

El motivo por el cual el exponente  $\gamma$  efectivo de la población es tanto mayor, es la distribución de la población en  $h$ . A pesar de que el valor medio de la fracción de energía heredada está en  $\langle h \rangle \approx 0.2$ , una gran parte de la población se compone de individuos con  $h \approx 0$ , como se muestra en la Figura 3.6. Se observó que esta distribución llega a un estado aparentemente estacionario. A tiempos suficientemente largos, las familias de  $\gamma$  mayor se reproducirían más rápido, tendiendo a una distribución de población tipo  $\delta(h)$ . Este régimen no alcanza a ser observado en nuestras simulaciones. El hecho de que los genes de la descendencia estén distribuidos uniformemente en la población hace que la independencia entre familias se rompa. Los miembros que nacen con  $h \approx 0$  tienen mayor probabilidad de sobrevivir y reproducirse, pero ya no aportan a su propio grupo genético sino que lo hacen a toda la población.

Dado que la distribución de la Figura 3.6 está normalizada, no da información acerca de la población total de cada una de las familias. Para esto, mostramos en la Figura 3.7 la evolución de las poblaciones medidas de una simulación con fracción de energía heredada variable. Se observa, al igual que para

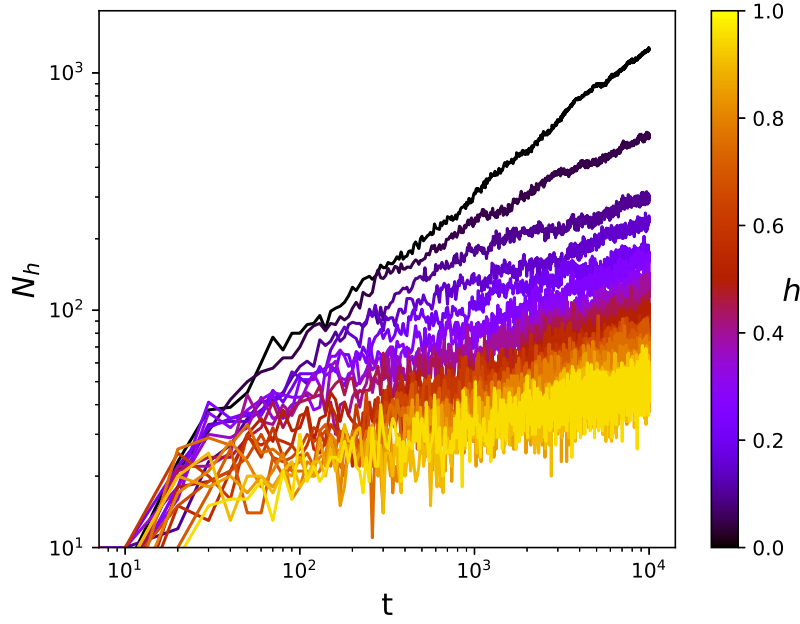


**Figura 3.5:** Exponentes  $\gamma$  para distintas fracciones de energía heredada  $h$  constantes. Se observa que es aproximadamente constante para  $0.2 < h < 0.8$ . En los extremos cambia abrupta pero continuamente. Además, para la simulación con fracción de energía heredada variable, el exponente es mucho mayor a la tasa correspondiente a una fracción constante (punto rojo).

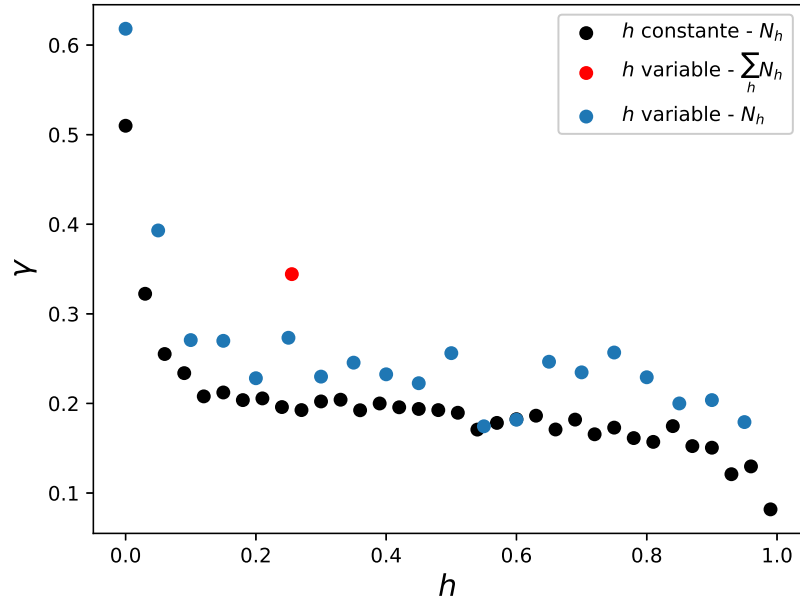


**Figura 3.6:** Distribución de la población en función de la fracción de energía heredada para distintos tiempos  $t$  (indicados por la escala de colores). Notamos que la población alcanza un estacionario que cumple empíricamente con la forma  $h^{-1}$ . La distribución converge porque para valores de  $h \approx 0$ , la distribución se aparta de la forma algebraica.

el caso constante, que las poblaciones crecen con una relación algebraica en el tiempo, por lo cual se puede definir un exponente  $\gamma$ , como se hizo con los datos de la Figura 3.1. En la Figura 3.8 se muestran estos exponentes, obtenidos a partir de la Figura 3.7 y se comparan con las simulaciones con fracción de energía heredada constante mostradas en la Figura 3.5. Se observa que, en todos los casos, la tasa de crecimiento es mayor para el caso variable que para el constante.



**Figura 3.7:** Dinámicas de las poblaciones correspondientes a distintos valores de fracción de energía heredada. Las familias evolucionan según una ley algebraica en el tiempo,  $N_h \propto t^{\gamma(h)}$  de la misma manera que las mostradas en la Figura 3.1.



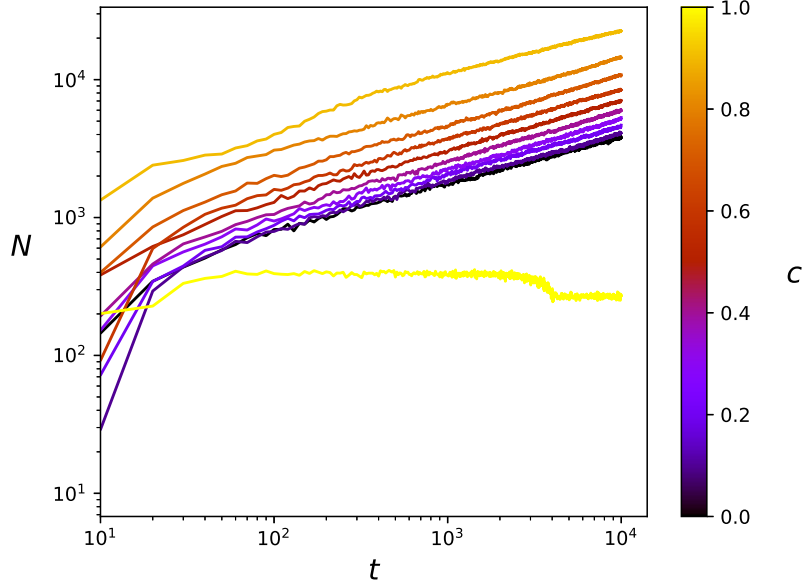
**Figura 3.8:** Exponentes de la población con fracción de energía heredada constante y variable. Se observa que en el segundo caso los exponentes son siempre mayores que en el primero.

## Con transmisión de información

A continuación, estudiamos cómo cambian los resultados anteriores al incorporar la transmisión de información genética. En la Figura 3.9 mostramos la evolución de la población total para distintas calidades de copia constantes  $0 \leq c \leq 1$ . La fracción de energía heredada  $h$  se dejó libre. Para una calidad de copia  $c = 0$ , la evolución coincide con la de la Figura 3.4, como es de esperarse. A medida que aumenta la calidad de copia, la población total a cada tiempo aumenta, conservando el comportamiento algebraico y el exponente  $\gamma$ . En  $c = 1$ , el sistema se somete a una transición de fase en la cual el comportamiento cambia radicalmente. Esto se debe a que, para  $c = 1$ , el sistema está



sujeto a la distribución genética inicial. Para evolucionar, no puede reemplazar gradualmente a unos genes por otros, probando nuevos genotipos de manera suave, sino que el linaje más apto de los que existían al principio extingue rápidamente al resto y luego se mantiene constante.

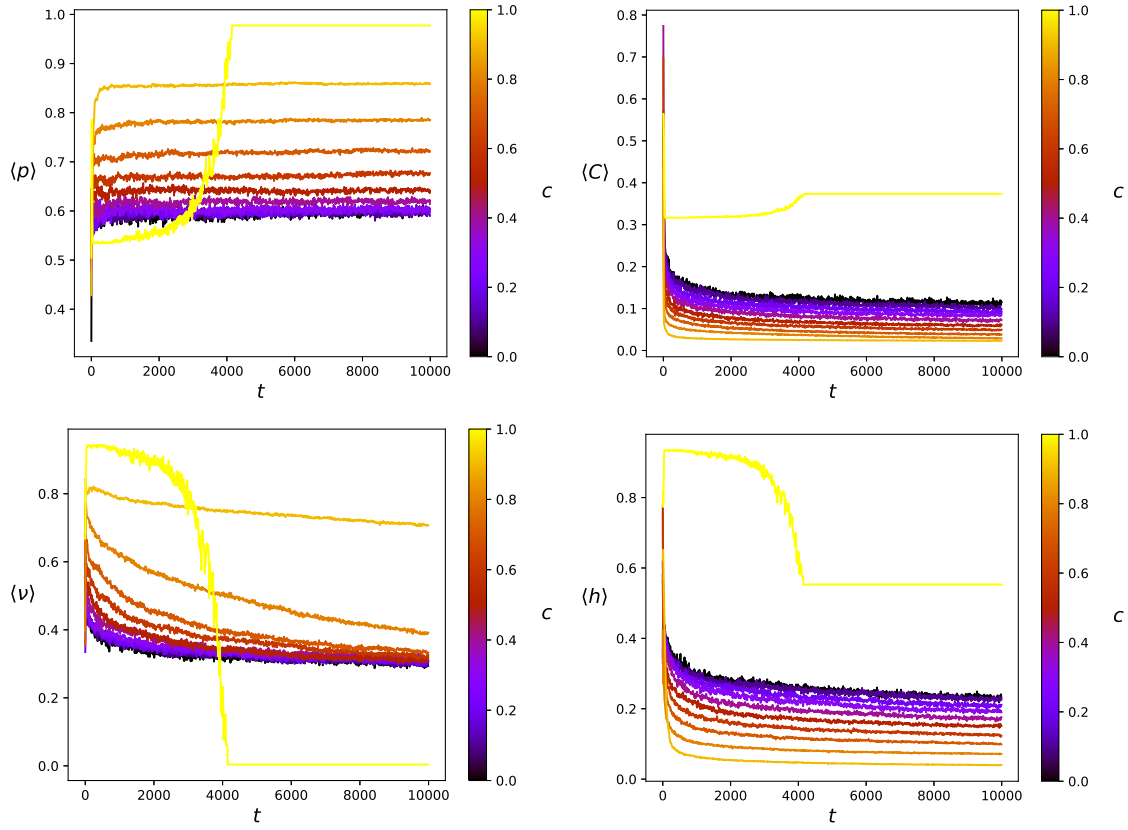


**Figura 3.9:** Evolución de la población para distintos valores constantes de la calidad de copia  $c$  para una fracción de energía heredada variable, al igual que el resto de los genes. Se observa que, a medida que aumenta la calidad de copia, el sistema conserva su comportamiento algebraico y el exponente  $\gamma$  pero con poblaciones cada vez mayores. Este comportamiento cambia en  $c = 1$ , donde el sistema presenta un cambio de fase debido a la imposibilidad de mutar de los genes con los que la simulación comenzó.

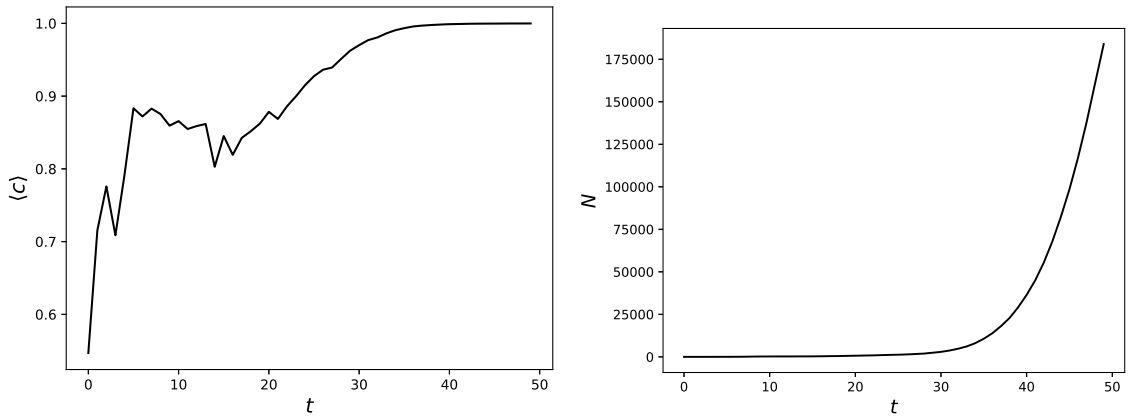
En la Figura 3.10 mostramos la evolución de los valores medios de los genes variables, es decir el pedido de recurso  $p$ , el consumo energético  $C$ , la tasa de reproducción  $\nu$  y la fracción de energía heredada  $h$ . Los comportamientos que más difieren de los hallados para calidad de copia  $c = 0$  son en el pedido de recurso y la tasa de nacimientos. En el primero, el hecho de tener alguna “seguridad” de que los genes van a ser similares elimina la neutralidad de este gen, agregándole una presión evolutiva hacia valores mayores. Para la tasa de reproducción se sigue observando un comportamiento decreciente, pero a medida que aumenta la calidad de copia, este decrecimiento se hace cada vez más lento. Una calidad de copia alta también permite que el consumo disminuya más rápido, asegurando que los descendientes también tengan un consumo bajo (si la calidad de copia es baja, los consumos de la descendencia estarán mucho más dispersos). A pesar de que la calidad de copia da “seguridad” sobre los genes de los descendientes, el sistema evoluciona a no entregarles energía ( $h \rightarrow 0$ ) ya que puede, en lugar de esto, asegurarles un pedido de recurso alto y consumo energético bajo, que es más beneficioso para los progenitores. Por otro lado, los valores estacionarios de la simulación con calidad de copia  $c = 1$  no son repetibles: en cada realización, la distribución genética inicial será distinta y, por lo tanto, el linaje más apto (que extinguirá a todos los demás y cuyos valores genéticos tomará la población en el estacionario) también lo será.

En la Figura 3.11 mostramos la evolución de la calidad de copia media  $\langle c \rangle$  y la población total  $N$  cuando la calidad de copia puede mutar con las mismas reglas que los otros genes. Se observa que este cambio produce un gran cambio en el comportamiento del sistema, con la población unos tres órdenes de magnitud por encima de lo que observamos con  $c = 0.9$ . La curvatura positiva al comienzo indica que la población empieza con un exponente  $\gamma > 1$  pero decrece a medida que  $\langle c \rangle$  se estabiliza cerca de 1.

Finalmente, en la Figura 3.12 mostramos la energía media por individuo de la población en función



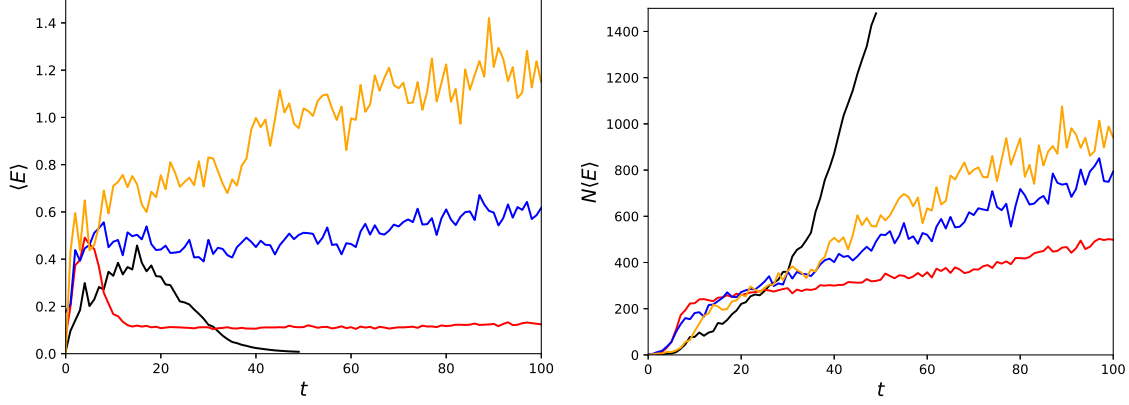
**Figura 3.10:** Evolución de los valores medios de los genes variables. Se observa que el pedido de recurso pierde su carácter de gen neutro, recibiendo mayor presión evolutiva a medida que aumenta la calidad de copia. La tasa de reproducción, por otra parte, pierde mucha de la presión debido a una nueva “confianza” en la descendencia. A pesar de esta, la fracción de energía heredada tiende a disminuir, con el sistema prefiriendo achicar el consumo energético más eficientemente y no perjudicar a los progenitores. Los valores estacionarios para calidad de copia  $c = 1$  no deben tomarse como referencia ya que estos dependen de los valores iniciales del sistema.



**Figura 3.11:** Calidad de copia media  $\langle c \rangle$  y población  $N$  en función del tiempo. Se observa que el hecho de hacer variable la calidad de copia hace que la población pueda crecer mucho más rápidamente. A medida que su valor se acerca cada vez más a uno, el exponente de la población se hará cada vez más lento (comienza mayor a uno y terminará igual que el de las poblaciones con calidad de copia constante de la Figura 3.9). Sin embargo, no podemos mostrar este régimen debido a la gran capacidad computacional requerida.

del tiempo para algunos valores de calidad de copia constante y para la calidad de copia variable. Observamos que a medida que aumenta la calidad de copia, disminuye la energía media por individuo. Esto podría interpretarse como un aumento de la eficiencia del sistema, pero debe tenerse en cuenta que esto no es contradictorio con lo discutido acerca de la disminución de la entropía y degradación

de la energía. La diferencia está en que la energía total de la población, que mostramos en la misma figura, aumenta con la calidad de copia. A pesar de que cada uno de los individuos es más eficiente, esto permite que haya más individuos y que la población gaste una mayor cantidad de energía.



**Figura 3.12:** Evolución de la energía almacenada por individuo, que disminuye con el aumento de la calidad de copia. Esto sugiere que los individuos son más eficientes. Sin embargo, la energía total almacenada por la población, también mostrada, aumenta con la calidad de copia. Esto indica que la población total es menos eficiente a pesar de que sus individuos no lo son. El hecho de que la población evoluciona hacia una calidad de copia mayor (y por lo tanto menor eficiencia grupal) refuerza lo discutido acerca de la entropía y degradación de la energía.

### 3.2. Información interna

En esta sección describimos una idea que quisiéramos implementar en las simulaciones, pero no alcanzamos a obtener resultados para agregar al trabajo. Sin embargo, consideramos que la idea puede ser de interés para trabajos futuros, o incluso como una interpretación acerca de como este tipo de modelo puede generar inteligencia artificial.

En la Sección 2.2 describimos el problema de tener un recurso con una dinámica de recuperación continua e individuos sin un mecanismo de control de reproducción. Una posible forma de implementar este tipo de controles es que la tasa de reproducción sea una combinación lineal de propiedades externas (en principio solo el recurso, pero podría ser también el número de individuos, por ejemplo) pero también internas. Así,  $\nu(t+1) = \sigma(\sum a_i \mathbf{F}_i(t) + \beta N(t) + \gamma f(t))$ , donde  $\sigma$  es una función normalizadora. Sin embargo, si lo hacemos para la tasa de reproducción, ¿por qué no hacerlo con los otros elementos del fenotipo? En general, podemos plantear que  $\mathbf{F}(t) = \sigma(\mathcal{A}\mathbf{F} + \beta f(t) + \gamma N(t))$ , donde  $\mathcal{A}$  es una especie de matriz de correlación de los elementos del fenotipo.

Ahora también podemos preguntarnos si la matriz  $\mathcal{A}$  y los vectores  $\beta$  y  $\gamma$  son elementos del genotipo o del fenotipo. Si son parte del fenotipo, podemos usar una matriz  $\mathcal{A}'$  para determinar los valores de  $\mathcal{A}$  de la misma manera que se hizo con  $\mathbf{F}$ . Esto constituye una segunda capa en una especie de red neuronal que determina, a partir del “estado sensorial” (en este ejemplo usamos  $\mathbf{F}$ ,  $f$ , y  $N$  como el estado sensorial) el estado fenotípico al tiempo siguiente. Para determinar completamente la red, faltan dos cosas: su estructura y el método de aprendizaje. Sugerimos que ambos compongan el genotipo de los individuos. De una generación a la otra, pequeñas fluctuaciones en la estructura y un funcional sobre la energía a cada tiempo (el método de aprendizaje) irá seleccionando naturalmente las mejores estructuras y técnicas.

El método propuesto para obtener comportamientos inteligentes puede ser bastante complicado de lograr. Con modelos más sencillos, como los que mostraremos en los Capítulos 4 y 5 mostraremos que también surgen comportamientos no triviales dignos de ser destacados sin la necesidad de emplear

redes neuronales.

## Capítulo 4

# Modelo de la hidra inmortal

En este capítulo presentaremos el análisis y resultados de un modelo basado en las características fisiológicas de los seres vivos macroscópicos. Dicho modelo es el que dio lugar a mi tesis de licenciatura [45]. Lo incluimos aquí porque es el primer modelo basado en energía que estudiamos, por lo cual puede servir de referencia para entender los orígenes de las ideas en las que se basa este trabajo. Además, establece una conexión con modelos de poblaciones comúnmente utilizados en la matemática de los sistemas biológicos [46], la cual se profundiza en la versión mejorada del modelo que se presenta en el Capítulo 5.

A diferencia de los modelos discutidos en capítulos anteriores, vamos a considerar aquí un sistema espacialmente extendido. Consideremos una población de individuos que ocupa un recinto de  $L \times L$  celdas con condiciones de borde reflectantes. Cada celda contiene un recurso que, en ausencia de consumo, crece con una dinámica logística de tasa de recuperación  $r$  y capacidad de carga  $K$ . En general,  $r$  y  $K$  podrían depender de la posición, pero supondremos que son espacialmente uniformes. Así, la celda en la posición  $\mathbf{x}$  crecerá según:

$$R(f(\mathbf{x})) = rf(\mathbf{x}) \left(1 - \frac{f(\mathbf{x})}{K}\right). \quad (4.1)$$

La dinámica logística produce un crecimiento exponencial (con tasa de crecimiento  $r$ ) para pequeñas cantidades del recurso  $f$  pero satura su crecimiento en la capacidad de carga  $K$ . Consideramos que es el modelo más simple para describir el comportamiento de un recurso real en términos energéticos ya que, agregando una corrección de segundo orden, evita el crecimiento ilimitado que supondría una energía ilimitada. Además, se conocen diversos ejemplos para los cuales la dinámica logística provee una descripción muy precisa de la dinámica poblacional [47]. Más detalles de la dinámica logística se dan en la Sección 4.1.

Los miembros de la población tienen dos propiedades macroscópicas codificadas en su genoma  $\mathbf{G}$ : su tamaño físico, al que nos referiremos indistintamente como masa, y su velocidad de crecimiento  $s$ ; es decir  $\mathbf{G} = (M, s)$ . El fenotipo  $\mathbf{F}$  está compuesto por la masa instantánea  $m$  que determina el máximo de energía,  $E = m$ , que el individuo puede almacenar y que evoluciona de acuerdo a una logística:

$$\frac{dm}{dt} = sm \left(1 - \frac{m}{M}\right), \quad (4.2)$$

un tamaño de paso  $l = m^{1/3}$ , que determina la cantidad de celdas que se desplaza un individuo a cada paso de tiempo, y un contador  $c$  que indica la cantidad de pasos consecutivos durante los cuales la densidad de energía,  $\varepsilon = E/m$ , se ha mantenido por encima de un umbral ( $\varepsilon_U = 0.8$  para individuos adultos, con  $m \geq 0.95M$ ). Este último rasgo es el que hace la diferencia del caso estudiado en la

Sección 2.2. El fenotipo es, entonces  $\mathbf{F} = (m, l, c, \mathbf{x})$ . Si el contador alcanza una cuenta de reproducción  $c_R = 400$ , el individuo entra en gestación, aumentando la masa estacionaria de su crecimiento al doble ( $M \rightarrow 2M$  en la Ecuación (4.2)). Este aumento de masa llega hasta  $m = 1.05M$ , momento en el que se separa la hija (con masa  $m = 0.05M$ ) de la madre (que queda con masa  $M$ ). El genoma de la hija es  $\mathbf{G}' = (M', s')$ , donde cada una de las componentes difiere hasta en un 5 % de las correspondientes de  $\mathbf{G}$ , con distribución uniforme.

El fenotipo del individuo incurre en un costo energético  $C$  que puede separarse en tres términos. El primero es el costo metabólico  $C_{met}$  por sostener con vida la masa  $m$ , y obedece a una relación alométrica con exponente  $3/4$  [48]. El segundo es el costo cinético  $C_{cin}$  por mover una masa  $m$  a velocidad  $l$  por unidad de tiempo (dado por la energía cinética). Finalmente, contamos el costo de crecimiento  $C_{crec}$  desde la masa de nacimiento hasta la adulta. La forma funcional de estos términos está dada por

$$\begin{aligned} C_{met} &= \mu m^{3/4}, \\ C_{cin} &= \kappa m l^2 = \kappa m^{5/3}, \\ C_{crec} &= m s \left(1 - \frac{m}{M}\right), \end{aligned} \quad (4.3)$$

donde los parámetros  $\mu$  y  $\kappa$  son constantes e iguales para toda la población a todo tiempo, y tienen poco efecto en la dinámica de la simulación.

Como mencionamos antes, la cantidad de energía que un individuo puede almacenar en su cuerpo está dada por su masa, con lo cual su consumo de recurso está limitado ya sea por la disponibilidad de lugar en la reserva, o la disponibilidad del recurso mismo en el sitio en el que el individuo se encuentra, ya que la alimentación es estrictamente local. Para tener en cuenta ambos factores, la función de consumo se puede escribir como:

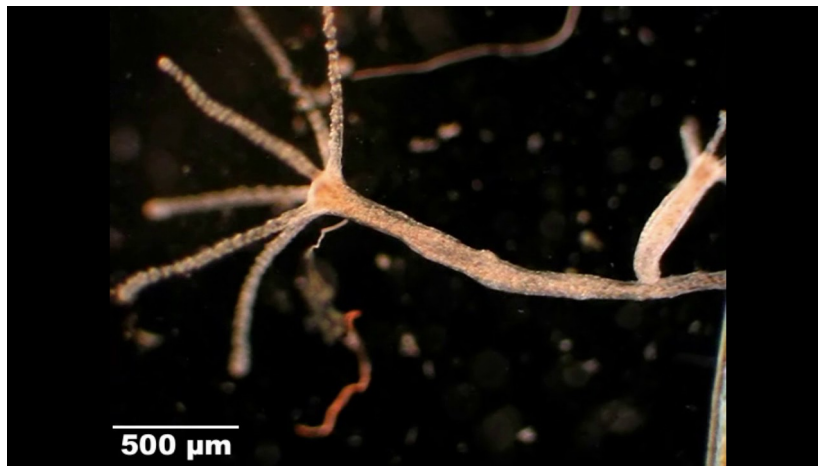
$$\varphi(\mathbf{F}, f) = \min(m - E, f(\mathbf{x}) - K\beta), \quad (4.4)$$

donde  $\beta = 0.01$  sirve para que el recurso en una posición no recaiga al equilibrio (inestable) de  $f(\mathbf{x}) = 0$ . La Ecuación (4.4) puede interpretarse como que los individuos ingieren lo suficiente como para llenarse si el alimento alcanza, o lo que esté disponible en caso contrario. Así, la dinámica energética total puede resumirse en:

$$\frac{dE}{dt} = \min(m - E, f(\mathbf{x}) - K\beta) - C_{met} - C_{cin} - C_{crec}. \quad (4.5)$$

La muerte de estos individuos ocurre solo cuando la dinámica energética total lleva la energía a cero, es decir que son biológicamente inmortales. Este modelo puede parecer un tanto artificial para describir organismos macroscópicos, especialmente por la reproducción asexual y la inmortalidad biológica. Sin embargo la *Hydra vulgaris*, que mostramos en la Figura 4.1, tiene características fenomenológicamente similares a las del presente modelo. Es un organismo invertebrado de agua dulce cuya locomoción se basa en dar tumbos, se reproduce asexualmente y es biológicamente inmortal. Por esta razón la elegimos para dar nombre al modelo.

A continuación mostraremos los resultados que arroja el modelo de la hidra inmortal. Queremos recordar que, a pesar de la similitud mencionada con el organismo del mismo nombre, el objetivo de este capítulo no es predecir o explicar propiedades de *Hydra spp.*, sino estudiar las propiedades del sistema, sujeto a selección natural como establecimos en el Capítulo 1.



**Figura 4.1:** Fotografía de una hidra. Del lado derecho se observa una hidra en gestación, que eventualmente crecerá y se separará de la madre, transformándose en un organismo independiente.

## 4.1. Dinámica poblacional

En esta sección estudiaremos la dinámica poblacional, es decir la evolución temporal del número de individuos vivos  $N(t)$ , obtenida con el modelo de la hidra inmortal. Las poblaciones iniciales para las simulaciones fueron de 1000 individuos repartidos uniformemente en el espacio genético, es decir  $\mathbf{G} = (M, s) \in (0, 1000] \times [0, 1)$ .

A continuación compararemos los resultados con la dinámica poblacional predicha por otros modelos que consideran al recurso disponible para la población en distintos regímenes.

### Modelo de Malthus

El primer modelo que consideramos es el crecimiento malthusiano [18]. Este predice una dinámica exponencial en la población dada por la ecuación

$$\frac{dN}{dt} = \alpha N, \quad (4.6)$$

donde  $\alpha$  es la diferencia entre las tasas de natalidad y mortalidad de la población. Debido a que ambas (y por tanto, también su diferencia) son constantes, la población diverge o se extingue. Las extinciones están asociadas con una población donde no ocurren nacimientos y las muertes ocurren con una tasa fija y por lo tanto no aportan mucho a este trabajo. Las divergencias, por otra parte, requieren de recurso ilimitado y queremos verificar si esta dinámica emerge del modelo sin incluirla explícitamente, por más de que no sea una situación realista.

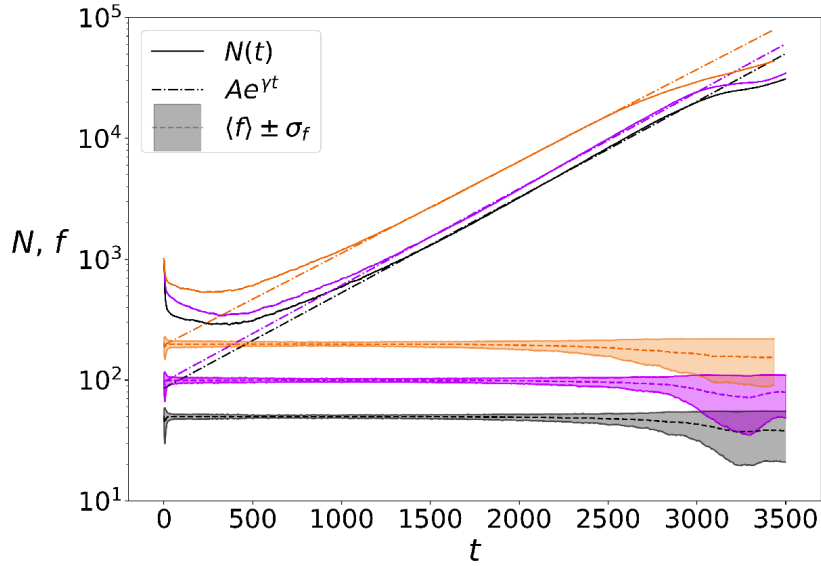
Para lograrlo, las simulaciones se inician con una capacidad de carga muy grande,  $50 \leq K \leq 200$  y todos los sitios con recurso completo, y una tasa de recuperación baja,  $r = 0.01$ . Así podrá observarse el régimen exponencial y la transición de recurso infinito a limitado ocurrirá lentamente.

En la Figura 4.2 se muestran los resultados de la simulación para tres valores de  $K$ . Se observan tres regiones principales en la evolución. La primera corresponde a un período de extinción donde los genes que consumen más que lo que retribuyen, desaparecen. Luego de esta, las poblaciones adoptan una dinámica exponencial,  $N(t) = A \exp(\gamma t)$ . Las líneas rectas (notar la escala vertical logarítmica) representan los mejores ajustes al crecimiento poblacional en esta región. El coeficiente de proporcionalidad  $A$  contiene información sobre el grado de extinción producida al principio de la simulación ya que representa la población inicial “aparente” en el régimen exponencial. Como la tasa de muertes es nula en esta región, el período de duplicación  $\tau$  (tiempo que demora la población en duplicar su tama-

ño) está relacionado con el coeficiente  $\gamma$  a través de la relación  $\gamma = \ln 2/\tau$ . El período de duplicación, a su vez, es el tiempo entre dos reproducciones sucesivas, y coincide con las cuentas de reproducción  $c_R$  cuando haya recurso disponible. Así,

$$\gamma = \frac{\ln 2}{c_R} \approx 0.00173, \quad (4.7)$$

que coincide con los valores de aproximadamente  $\gamma = 0.0018$  obtenidos en el ajuste. La última región muestra que la dinámica poblacional se aparta gradualmente de la exponencial. En esta región también se observa que el valor medio del recurso, marcado por la línea aproximadamente horizontal, comienza a disminuir, mostrando que el recurso deja de comportarse como inagotable. Antes de esto, el consumo de la población tenía efecto solo sobre la desviación estándar del recurso, marcada por la región sombreada alrededor del valor medio.



**Figura 4.2:** Evolución temporal de la población para capacidades de carga  $K = 50$  (negro),  $K = 100$  (magenta) y  $K = 200$  (naranja). Se observa una extinción inicial que no afecta el valor medio del recurso disponible, por lo cual la atribuimos a genes que no pueden sobrevivir en ninguna situación debido a su dinámica energética. Luego hay un régimen exponencial con exponente  $\gamma = 0.0018$  para las tres curvas, en concordancia con lo esperado por el período de evaluación del entorno que impone la dinámica de reproducción. En este régimen el valor medio del recurso también se mantiene constante, demostrando que efectivamente se comporta como inagotable. Finalmente, la evolución se aparta del comportamiento exponencial al mismo tiempo que el valor medio del recurso empieza a cambiar.

## Modelo logístico

El modelo logístico [19] es una primera corrección al modelo de Malthus para tener en cuenta que los recursos disponibles para mantener viva a una población son finitos. Consiste en agregar un término negativo de segundo orden en  $N$  en la dinámica energética. Esta puede interpretarse como que la tasa de muertes de la población aumenta con el número de individuos porque el recurso se reparte entre ellos (competencia intraespecífica) [47]. Para una tasa de recuperación  $r'$  y capacidad de carga  $K'$ , la forma habitual de escribir la ecuación logística es

$$\frac{dN}{dt} = r'N \left( 1 - \frac{N}{K'} \right), \quad (4.8)$$



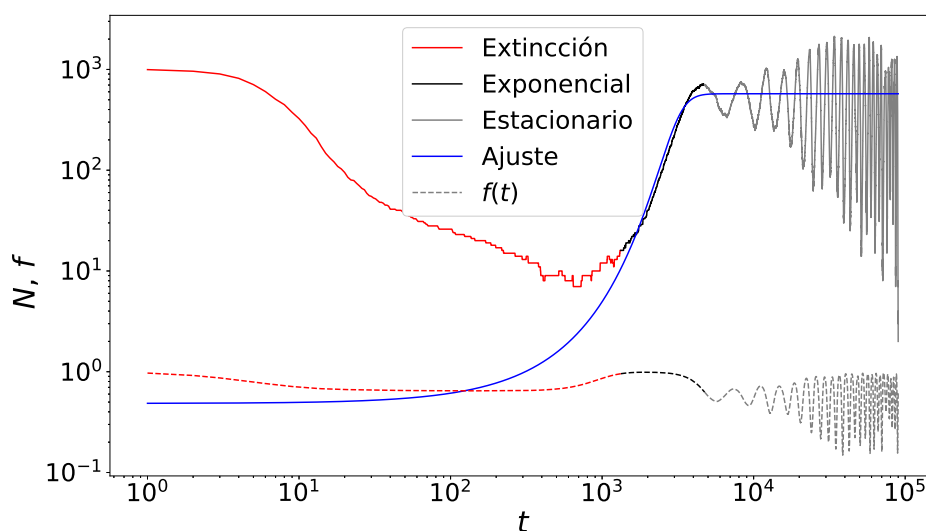
que, para una condición inicial  $N(0) = N_o$  tiene por solución

$$N(t) = \frac{K'}{1 + \frac{K' - N_o}{N_o} \exp(-r't)} \quad (4.9)$$

Repitiendo simulaciones similares a las de la Sección 4.1, pero con la capacidad de carga del recurso en  $K = 1$ , puede observarse el régimen transitorio que se asemeja a la evolución exponencial de la misma sección y la evolución en un régimen de recurso limitado. Como se observa en la Figura 4.3, este régimen es claramente oscilatorio mientras que la ecuación logística predice un equilibrio asintóticamente estable en  $K'$  (ya que  $dR(f)/df < 0$  en  $K'$ ). Se sabe que para el caso discreto de la ecuación logística, realizaciones individuales producen oscilaciones mientras que el valor medio no [49]. Sin embargo, a diferencia del caso de nuestras simulaciones, las oscilaciones de la ecuación logística discreta mantienen su amplitud. De todos modos, ajustar los parámetros de la Ecuación (4.9) a las curvas de población (ignorando hasta el primer mínimo) produce dos resultados relevantes. Primero, permite definir sistemáticamente el régimen de extinción, el de crecimiento exponencial, y el estacionario (aunque sea oscilatorio) para tener un método más riguroso para comparar resultados. Además, los residuos del ajuste proveen un método para cuantificar la diferencia entre el modelo logístico y los resultados numéricos.

En la Figura 4.3 se observan las tres regiones distinguidas por color, mostrando el recurso (línea a trazos) y la población (línea continua). El criterio tomado para elegir la zona exponencial a partir de los parámetros del ajuste está basado en la escala típica de variación de la logística en su punto de mayor derivada, que es también el punto de inflexión en el ajuste de la población. Evaluando la Ecuación (4.8) en  $(N_o + K')/2$ , la escala de tiempo típica del estacionario ( $\Delta N \approx 1$ ) es

$$\Delta t \approx 4 \frac{K' - N_o}{r' K'}. \quad (4.10)$$



**Figura 4.3:** Ajuste de la Ecuación (4.9) a los resultados de las simulaciones con capacidad de carga del recurso  $K = 1$ . Se observa claramente la diferencia entre el equilibrio constante dado por la ecuación logística (azul) y las oscilaciones observadas en las simulaciones (gris). También se observa la separación en regiones que se obtiene con este ajuste, eligiendo la región exponencial.

Con el criterio elegido, se observa que la región exponencial coincide con la noción de recurso infinito presentado en la sección anterior, ya que en ella el recurso permanece aproximadamente constante. Como se observa en el régimen “estacionario” de la dinámica de población, hay una sucesión de oscilaciones de amplitud creciente que sugieren que la dinámica logística no es suficiente para describir

el sistema.

## Modelo de Lotka-Volterra

El último modelo con el que queremos comparar la dinámica poblacional es el modelo de Lotka-Volterra [46]. El modelo de Lotka-Volterra considera las interacciones entre un depredador y su presa a través de interacciones de campo medio. Para una presa  $\mu$  y un depredador  $\nu$ , sus ecuaciones fundamentales son

$$\begin{aligned}\frac{d\mu}{dt} &= c_{11}\mu - c_{12}\nu\mu, \\ \frac{d\nu}{dt} &= c_{21}\mu\nu - c_{22}\nu.\end{aligned}\tag{4.11}$$

Las Ecuaciones (4.11) representan un sistema donde la presa crece ilimitadamente en ausencia de depredadores, mientras que estos decrecen exponencialmente en ausencia de presas. Está claro que una diferencia inmediata con el modelo de la hidra inmortal es que las presas (el recurso) crecen con una dinámica logística (no exponencial). Sin embargo, el comportamiento cualitativo del modelo sigue siendo de interés. Suponiendo que la población  $N$  son los depredadores y el recurso  $f$  las presas, podemos escribir las ecuaciones de la siguiente manera:

$$\begin{aligned}\frac{d \ln(\langle f \rangle)}{dt} &= c_{11} - c_{12}N, \\ \frac{d \ln(N)}{dt} &= c_{21}\langle f \rangle - c_{22}.\end{aligned}\tag{4.12}$$

De esta manera, los lados izquierdo y derecho de las igualdades quedan separados entre depredador y presa, y podemos ajustar los coeficientes  $c_{ij}$  con los resultados de las simulaciones. Nótese que si se considera la cantidad de recurso total disponible,  $\sum_{\mathbf{x}=1}^{L^2} f(\mathbf{x}) = L^2\langle f \rangle$ , el sistema de ecuaciones es equivalente a otro sistema de Lotka-Volterra ya que:

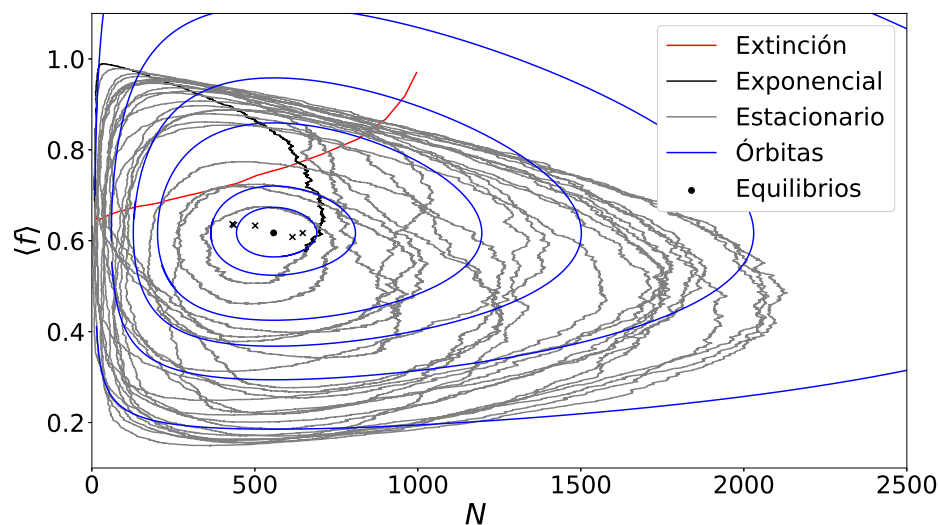
$$\begin{aligned}\frac{d \ln(L^2\langle f \rangle)}{dt} &= \frac{d \ln(\langle f \rangle)}{dt}, \\ c_{21} &\rightarrow L^2 c_{21}.\end{aligned}\tag{4.13}$$

En la Figura 4.4 se observa la trayectoria trazada por  $(N(t), \langle f \rangle(t))$  durante la evolución del sistema en el espacio de fases. A partir de los coeficientes  $c_{ij}$  ajustados puede calcularse el único equilibrio estable,  $(N, \langle f \rangle)_{eq} = (c_{22}/c_{21}, c_{11}/c_{12})$ . En la misma figura se muestran algunos equilibrios calculados a partir de trayectorias obtenidas en distintas realizaciones del mismo experimento. A diferencia de los sistemas de Lotka-Volterra, se observa que la trayectoria simulada no presenta órbitas cerradas. Además, el equilibrio calculado parece ser inestable (la trayectoria del espacio de fases se aleja de él).

En definitiva, el modelo de Lotka-Volterra es una mejora cualitativa respecto a los modelos anteriores, en el sentido de que incluye interacción entre el recurso y la población. Sin embargo, no es suficiente cuando se quiere tener en cuenta efectos de estocasticidad, los cuales están muy presentes en el modelo de la hidra inmortal.

## 4.2. Más allá de los modelos estándar

En esta sección presentaremos algunas propiedades adicionales del modelo de la hidra inmortal. Estas propiedades lo diferencian de otros modelos, dándole un valor propio en cuanto al entendimiento de los sistemas biológicos.



**Figura 4.4:** Trayectoria del espacio de fases  $(N, \langle f \rangle)$ , equilibrios obtenidos a partir del ajuste de los coeficientes  $c_{ij}$  y las órbitas correspondientes al ajuste de la trayectoria simulada. Esta última aumenta su amplitud, sugiriendo que los equilibrios calculados son inestables. Además, los equilibrios parecen mantener un valor constante en  $\langle f \rangle$  pero variar en  $N$ .

## Probabilidad de supervivencia

Una de las implicaciones más inmediatas de un sistema recurso-población es que, en ausencia de recurso, la población se extingue. El efecto Allee [50], por ejemplo, tiene en cuenta un umbral inferior de la población debajo del cual esta se extingue. En muchos sistemas naturales esta cota no necesariamente es estrictamente cero, principalmente por efectos de la distribución espacial del recurso y su consumo localizado [50].

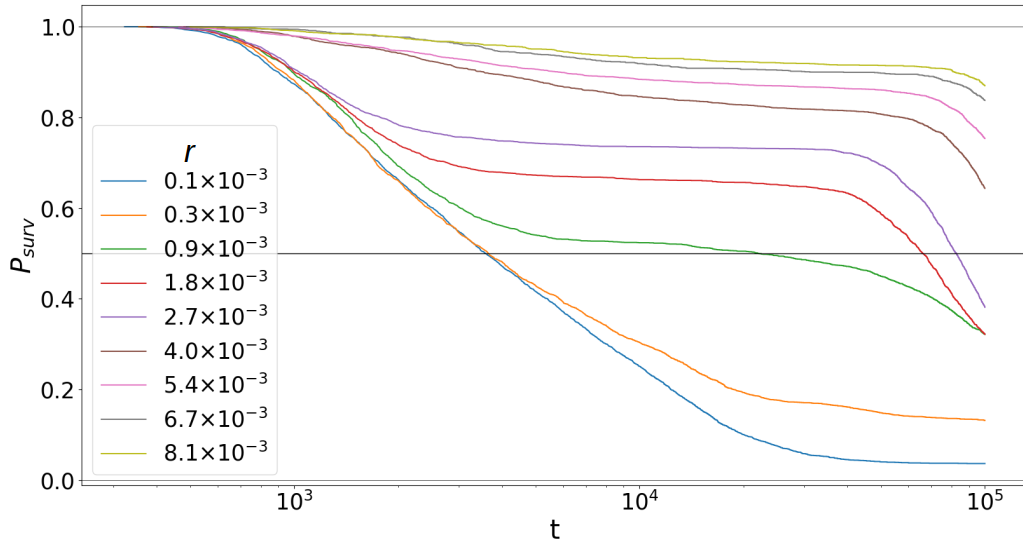
Para caracterizar los parámetros del recurso en los cuales la población sobrevive, hicimos cien realizaciones de una simulación como la mencionada anteriormente para nueve velocidades de recuperación, a capacidad de carga constante. Repetimos el experimento para nueve capacidades de carga a velocidad de recuperación constante. Esto permitió estimar la probabilidad de supervivencia hasta cierto tiempo a partir de la cantidad de realizaciones en las que hay una población con al menos un habitante.

En la Figura 4.5 se muestra la estimación de la probabilidad de sobrevivir dadas ciertas condiciones iniciales en función del tiempo. Como es de esperar, a medida que aumenta la velocidad de recuperación para un tiempo fijo, aumenta la probabilidad de sobrevivir. Para velocidades de recuperación  $r \geq 0.9 \times 10^{-3}$  se observan regiones donde la probabilidad de supervivencia es aproximadamente constante. En todos los casos se observa una caída brusca luego de esta meseta que parece ocurrir a la misma velocidad, pero a tiempos posteriores, a medida que aumenta la tasa de recuperación.

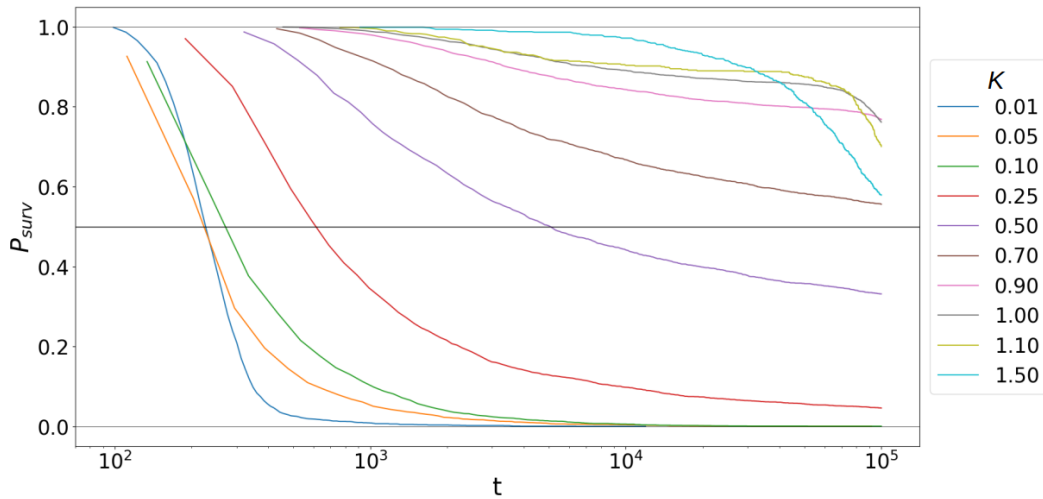
En la Figura 4.6 se muestra el análogo a la Figura 4.5, manteniendo constante la tasa de recuperación y variando la capacidad de carga  $0.01 \leq K < 2$ . Para capacidades de carga pequeñas, el comportamiento es el esperado: a mayor capacidad de carga, mayor probabilidad de sobrevivir a tiempo fijo. Sin embargo, si aumentamos más la capacidad de carga, empezamos a encontrar tiempos a los cuales la probabilidad de sobrevivir es mayor con una menor capacidad de carga.

## Oscilaciones de población

A pesar de que la amplitud de las oscilaciones de población y recurso no alcanzan un valor estacionario, comprobamos que el período  $T$  de las oscilaciones sí lo hace. Para caracterizarlo, se midieron los períodos de distintas realizaciones para algunos valores característicos de velocidad de recuperación



**Figura 4.5:** Probabilidad de supervivencia en función del tiempo para distintas tasas de recuperación  $10^{-4} \leq r < 9 \times 10^{-3}$ .

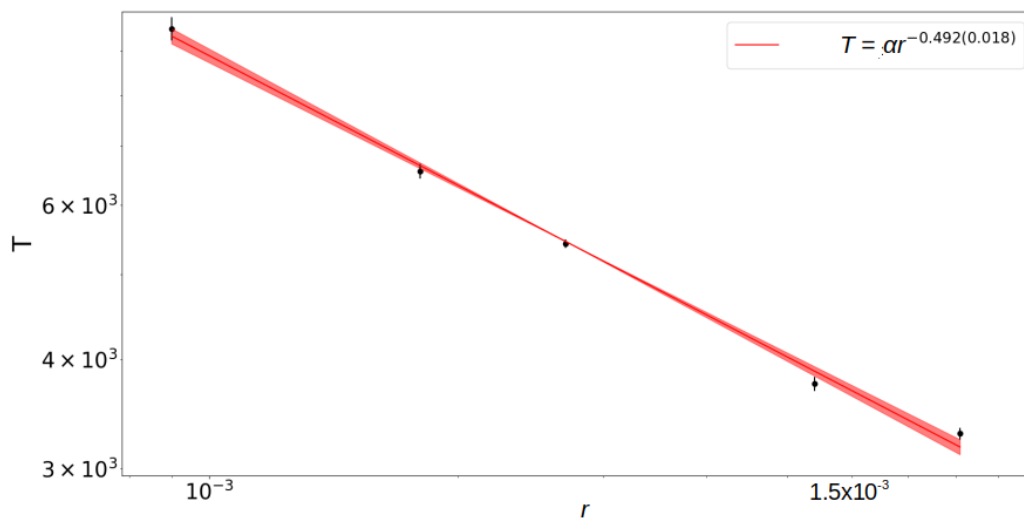


**Figura 4.6:** Probabilidad de supervivencia en función del tiempo para distintas capacidades de carga  $0.01 \leq K < 2$ .

(a capacidad de carga fija) y para algunos valores característicos de capacidad de carga (a tasa de recuperación fija).

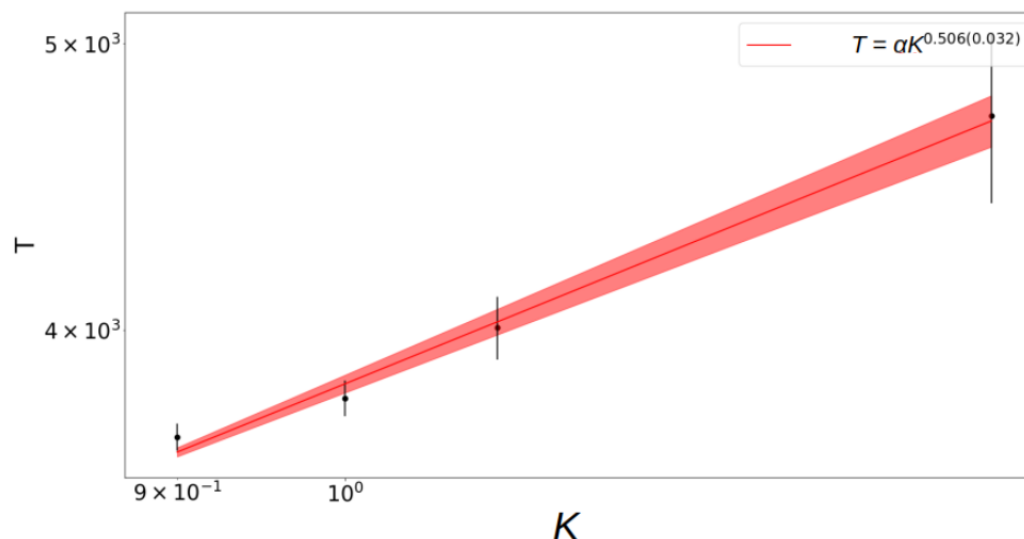
En la Figura 4.7 se muestran los períodos calculados para las distintas realizaciones a capacidad de carga constante, y un ajuste de la forma  $T = Ar^B$ . Como las poblaciones tienen más probabilidad de extinguirse a tasas de recuperación bajas, hay menos puntos para realizar los ajustes. Se observa que el período de las oscilaciones disminuye a medida que aumenta la tasa de recuperación. Como nacimientos y muertes consumen más energía (se consume el recurso en crecimiento o se priva a otros miembros de la población), pueden asociarse las oscilaciones de más frecuencia con mayor energía. Esta mayor energía proviene de la mayor tasa de recuperación. Es interesante notar que el exponente ajustado,  $B = -0.492$ , es muy cercano a  $-1/2$ . Sin embargo, no encontramos una explicación a esto.

En la Figura 4.8 se observan los períodos calculados para las distintas realizaciones a tasa de recuperación constante, y un ajuste de la forma  $T = A'K^{B'}$ . De la misma manera que con  $B$ , el exponente ajustado,  $B' = 0.506$ , sugiere una explicación analítica subyacente. Es interesante notar que la dispersión de los períodos es mucho mayor que en los experimentos a capacidad de carga



**Figura 4.7:** Período de las oscilaciones de población medido de las simulaciones y ajuste  $T = Ar^B$ . Destacamos el carácter decreciente del período con  $r$  y el hecho de que el coeficiente ajustado  $B$  tiene un valor muy cercano a  $-1/2$ , lo cual sugiere la posibilidad de encontrar una explicación analítica del coeficiente. Sin embargo, no pudimos encontrar tal explicación.

constante, sugiriendo que los períodos están mucho más controlados por la tasa de recuperación. Además, el período es creciente con la capacidad de carga. Con el mismo razonamiento que el del período para distintas tasas de recuperación, podemos asociar mayor capacidad de carga con menor energía. Esto puede parecer contraintuitivo, pero como la energía se consume en movimiento y mayor capacidad de carga permite quedarse en un lugar más tiempo, la asociación cobra sentido.



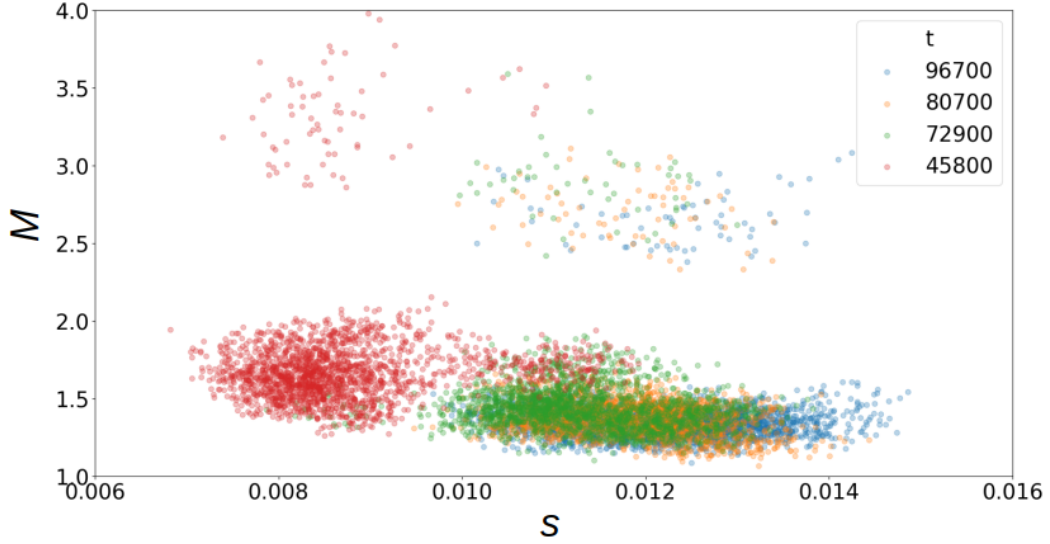
**Figura 4.8:** Período de las oscilaciones de población medido de las simulaciones y ajuste  $T = A'K^{B'}$ . Destacamos el carácter creciente del período con  $K$  y el hecho de que el coeficiente ajustado  $B'$  tiene un valor muy cercano a  $1/2$ , lo cual sugiere la posibilidad de encontrar una explicación analítica del coeficiente. Sin embargo, no pudimos encontrar tal explicación.

### 4.3. Dinámica genética

En la asociación de mayor capacidad de carga con mayor tiempo de permanencia, está implícito que los genes de la población son similares en ambos casos. De otra manera, la población podría

converger a una mayor masa que requiera más recurso y los tiempos de permanencia sean iguales. En esta sección justifiaremos esa suposición.

En la Figura 4.9 se observan las nubes de puntos que representan la distribución genética de la población en el espacio genético  $(s, M)$ . Graficamos cuatro nubes distintas que representan las distribuciones a distintos tiempos. Cada una de las nubes presenta otra más arriba, al doble de la masa, que corresponde a integrantes de la población que se encuentran gestando en ese momento.



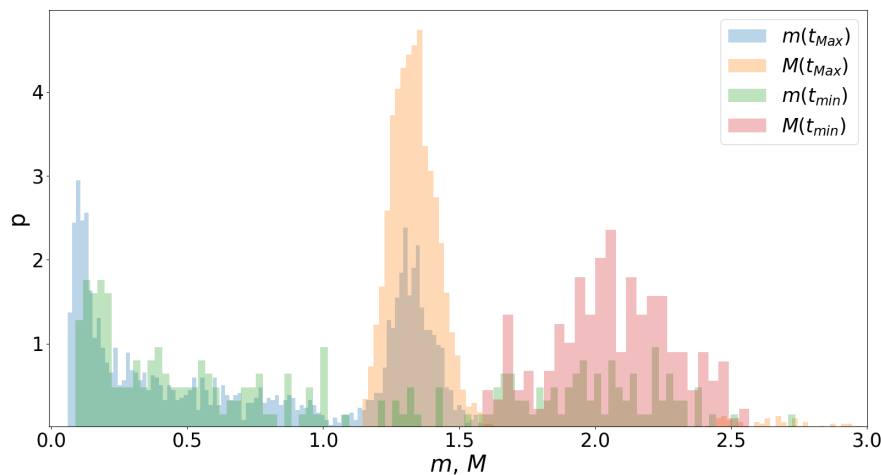
**Figura 4.9:** Nubes representativas de la población en el espacio genético para distintos tiempos. Se observa que la velocidad de crecimiento aumenta y la masa disminuye. Esto permite disminuir el costo cinético mediante el tamaño de paso, el metabólico y de crecimiento mediante la masa, y acorta el tiempo en llegar a la madurez mediante la velocidad de crecimiento.

Podemos observar dos tendencias principales respecto a la evolución genética. Dichas tendencias se repiten entre realizaciones y condiciones iniciales por lo cual presentamos una sola realización para mantener la claridad. La primera es que, al pasar el tiempo, la masa disminuye hacia el valor 1. Esto se entiende haciendo el siguiente razonamiento. La finalidad del movimiento es escapar de un sitio en el que se agotó el recurso. Además, un individuo no afecta los sitios vecinos mientras ocupa el central por lo que los sitios vecinos tienen, por lo general, igual probabilidad de tener alimento que uno más lejano. Como moverse más lejos cuesta más energía, pero no presenta beneficio, es perjudicial en neto y la presión evolutiva es hacia una masa unitaria. Si la masa es menor, el tamaño de paso es nulo (debido al carácter discreto de la grilla) y el individuo nunca se mueve de su sitio. Además, la masa no puede converger estrictamente a uno ya que el crecimiento logístico es asintótico hacia uno, y los individuos pasan la primera parte de su vida sin moverse.

El aumento de la velocidad de crecimiento es la segunda tendencia que se observa en la Figura 4.9. Como el costo energético de crecer es el aumento de masa, el tiempo en el cual el crecimiento ocurre no tiene efecto en el costo, y este puede crecer siempre que el individuo recupere la energía del crecimiento lo suficientemente rápido. Por el argumento referido al tamaño de paso convergiendo a un valor ligeramente superior a uno, a mayor velocidad de crecimiento la masa puede disminuir. Esto es conveniente ya que el costo energético del crecimiento es menor, el tiempo de crecimiento es menor, y los gastos cinéticos y metabólicos también disminuyen.

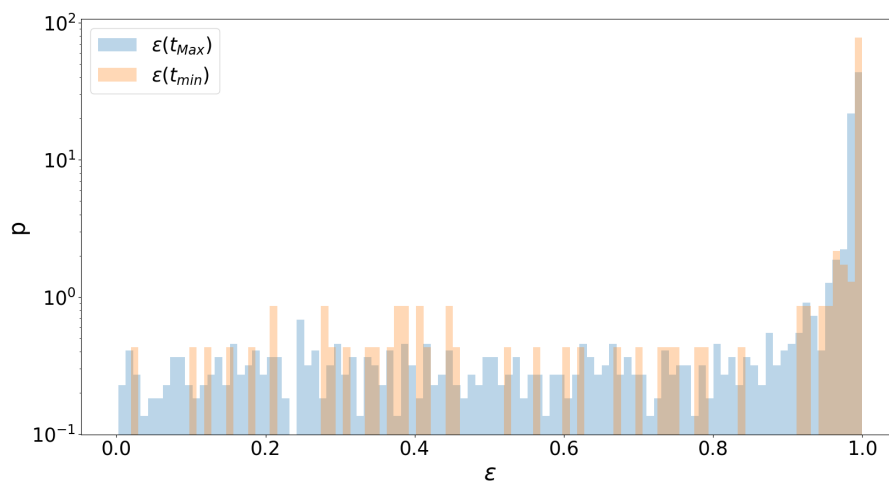
Las oscilaciones de población mencionadas en la sección anterior 4.2 indican que hay períodos de prosperidad ( $dN/dt > 0$ ) y de déficit ( $dN/dt < 0$ ). La Figura 4.10 muestra la distribución del gen de masa adulta y la masa de la población para un tiempo  $t_{Max}$  en el cual la población se encuentra en un máximo, y lo mismo para el tiempo  $t_{min}$  en el cual la población se encuentra en un mínimo. Se

observa que la seguridad que otorga un período de prosperidad hace a los genes relajarse hacia valores de menor consumo energético pero también más arriesgados.



**Figura 4.10:** Distribuciones del tamaño adulto y tamaño físico. Se muestran distribuciones al tiempo  $t_{Max}$  al cual la población se encuentra en un máximo y al tiempo  $t_{min}$  en el cual la población se encuentra en un mínimo. En el primero, la población está mucho más cerca de  $M \approx 1$ , mientras que en el segundo  $M \approx 2$ , debido a su mayor capacidad de almacenar energía.

También es interesante notar, como mostramos en la Figura 4.11, que la distribución de densidad de energía  $\epsilon = E/m$  en la población es muy similar cuando la población está en un máximo o un mínimo, con un pico muy marcado en  $\epsilon = 1$  y uniformemente distribuido para el resto del rango. Esto indica que a falta de recurso, la población se ve uniformemente afectada, sugiriendo que la causa de la falta es un exceso de población entre la cual se tiene que distribuir el recurso.



**Figura 4.11:** Distribución de energía en la población para tiempos  $t_{min}$  y  $t_{Max}$  de mínima y máxima población, respectivamente, en un período de las oscilaciones. Se observa que ambas tienen la misma forma por lo cual, dado que la distribución es una densidad, podemos afirmar que todos los individuos se ven afectados de la misma manera.

Algunos de los resultados de la evolución genética hallados en este capítulo pueden parecer contraintuitivos. En efecto, la evolución ha tendido hacia seres vivos que tienen un período de adaptación más largo (como es el caso de los mamíferos y, en particular, los seres humanos). Sin embargo, la diferencia fundamental con el modelo de la hidra es el aprendizaje. En la realidad, los seres vivos

alargan su período de dependencia a costa de poder aprender y realizar tareas más complejas. También se ha conjeturado que los individuos que sobreviven a períodos de falta de recurso (como las grandes extinciones) son los más pequeños [51], contrariamente a lo que sucede en las simulaciones. Sin embargo, el modelo de la hidra inmortal muestra que un sistema más complejo que los que se mostraron en capítulos anteriores también se somete a la selección natural, y en el próximo capítulo intentaremos mejorarlo.



## Capítulo 5

# Modelo de la hidra

*“Death is light as a feather; Duty, heavy as a mountain.”*

— al’Lan Mandragoran

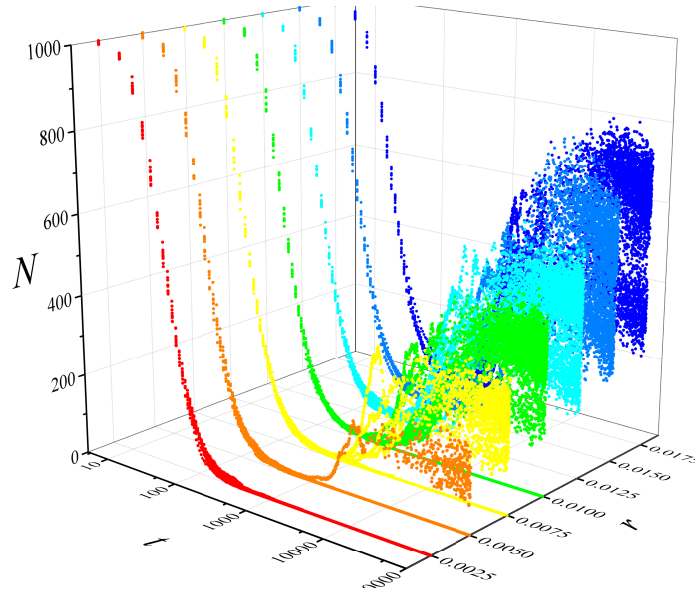
El modelo de la hidra es una versión revisada del modelo de la hidra inmortal presentado en el Capítulo 4. En este último se obtuvieron una serie de resultados interesantes, corroborando que la selección natural funciona con el modelo planteado y produce poblaciones persistentes en el tiempo. Sin embargo, muchas de las características genéticas de la población tienen evoluciones inesperadas o demasiado simples. Por ejemplo, no se observó especiación en el sentido de hallar grupos con características genéticas diferentes para las mismas condiciones iniciales.

El cambio principal que presentamos en este capítulo es en la dinámica de movimiento. En el modelo de la hidra inmortal, los individuos se mueven una distancia  $l$  en cada paso de tiempo. Como observamos que esto producía la convergencia de la población a un gen de masa  $m \approx 1$ , reemplazamos el tamaño de paso por una probabilidad  $p$  de desplazarse una celda en cada instante. Es decir, a cada paso de tiempo, el individuo se mueve con probabilidad  $p$  a uno de los sitios vecinos y se queda en el sitio con probabilidad  $1 - p$ . Además, se independizó esta probabilidad de la masa del individuo, agregando otro gen,  $P$ , que indica la probabilidad de moverse cuando el individuo tiene su masa igual a la masa adulta. La probabilidad de movimiento a cada edad evoluciona proporcionalmente a la masa instantánea, es decir  $p = Pm/M$  para capturar cualitativamente las posibilidades que da tener un cuerpo mayor. El segundo cambio, que inspira el cambio de nombre, es el agregado de una muerte por edad avanzada. Esta se implementó como un proceso de Poisson dependiente del tiempo y se describirá más adelante. Otros detalles menores, como el tiempo de evaluación de entorno  $c_R = 100$  y los valores de  $\mu$  y  $\kappa$  también se cambiaron.

Los resultados nuevos que hallamos tras estas modificaciones del modelo de la hidra fueron sorprendentes [52]. Encontramos que se produce especiación, que vale el principio de exclusión competitiva, y que existen comportamientos altruistas en un contexto de reglas egoístas y un comportamiento inteligente rudimentario, similar al de las especies reales.

### 5.1. Dinámica poblacional

En la Figura 5.1 se observan evoluciones temporales del número de individuos  $N$ . Se muestran diez realizaciones para cada tasa de recuperación  $r$ . El primer rasgo distintivo, sobre todo para tasas de recuperación grandes, es la aparición de dos valores de poblaciones estacionarias (notar la escala de tiempo logarítmica) que aumentan a medida que aumenta la tasa de recuperación. Se las puede



**Figura 5.1:** Evolución temporal del número de habitantes. Se observa, sobre todo para tasas de recuperación grandes, que las poblaciones estacionarias toman uno de dos valores estables.

apreciar como dos nubes de puntos, separadas por una brecha de menor densidad de puntos, en el régimen de tiempos largos.

A partir de la evolución temporal de la población podemos obtener el espectro de potencias de las fluctuaciones de población. En la Figura 5.2 se muestra este espectro, calculado como el módulo cuadrado de la transformada de Fourier de la población en el estado estacionario. Se pueden ver, además de los datos, ajustes de la forma  $f^{-\alpha}$ . Los coeficientes  $\alpha$  son indistinguibles entre sí dentro de la precisión del ajuste, y muy cercanos a  $\alpha = 1$ , sugiriendo una conexión con sistemas de criticalidad autoorganizada como los estudiados por Bak y Sneppen [53]. En la misma figura se observa un pequeño pico alrededor de  $f = 10^{-2}$ , señalado por una flecha ( $c_R^{-1}$ ). Este proviene del período de reproducción asintótico que coincide con el mínimo período de reproducción,  $c_R = 100$ . Presentaremos más detalles de esto en la próxima sección.

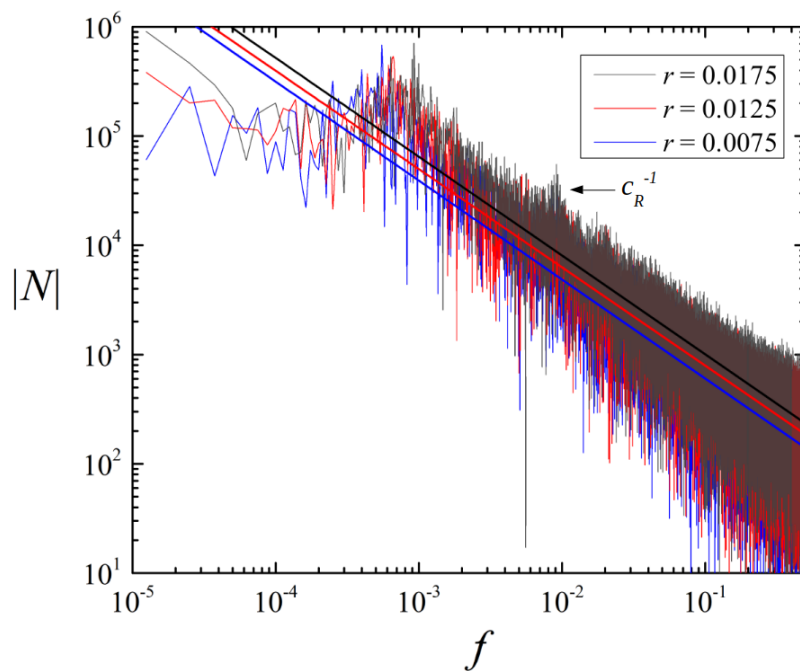
A través del teorema de Wiener-Khinchin [54] se puede vincular este espectro con la función de autocorrelación temporal de la población. Esto da una idea del tiempo de coherencia entre los individuos de la población. Los resultados se muestran en la Figura 5.3. Para los tres valores de tasa de recuperación mostradas, la autocorrelación empieza en un valor alto y, cerca del período correspondiente a la frecuencia asociada con el pico señalado en la Figura 5.2, cae repentinamente. Asociamos esta caída a la muerte de los individuos que, como mostraremos en la próxima sección, ocurre en este rango de edades de los individuos.

Ahora que hemos estudiado las características de la dinámica poblacional del sistema, presentaremos los resultados referidos a la dinámica genética de los individuos. Estos esclarecen algunas dudas surgidas en el estudio de la dinámica poblacional y a su vez plantean otras. Por esta razón, este análisis se constituye en el núcleo de esta tesis.

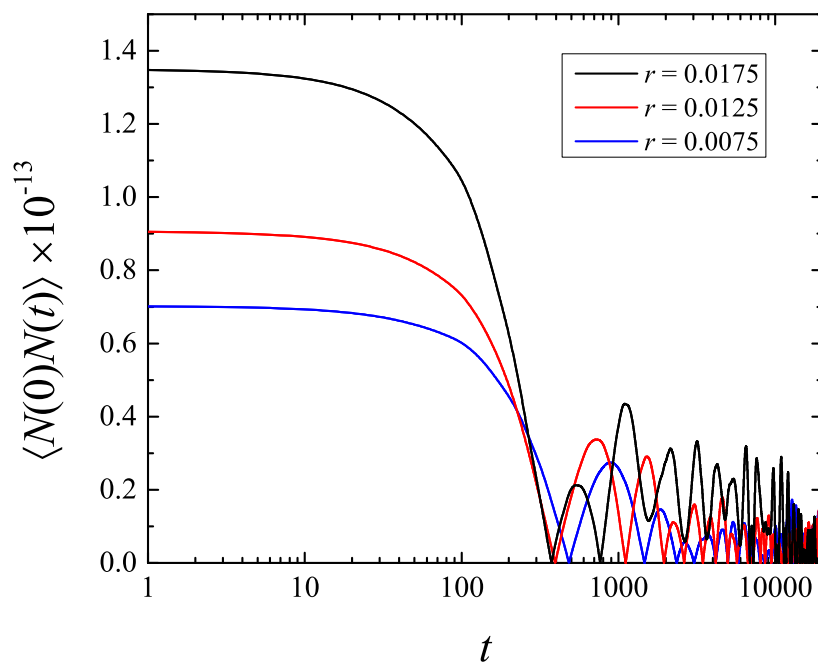
## 5.2. Dinámica individual

### 5.2.1. Genética

Como mostramos en la Figura 5.1, la cantidad de individuos converge a dos poblaciones posibles. Debido a que la cantidad de recurso disponible para ambas es la misma, esperamos que la mayor



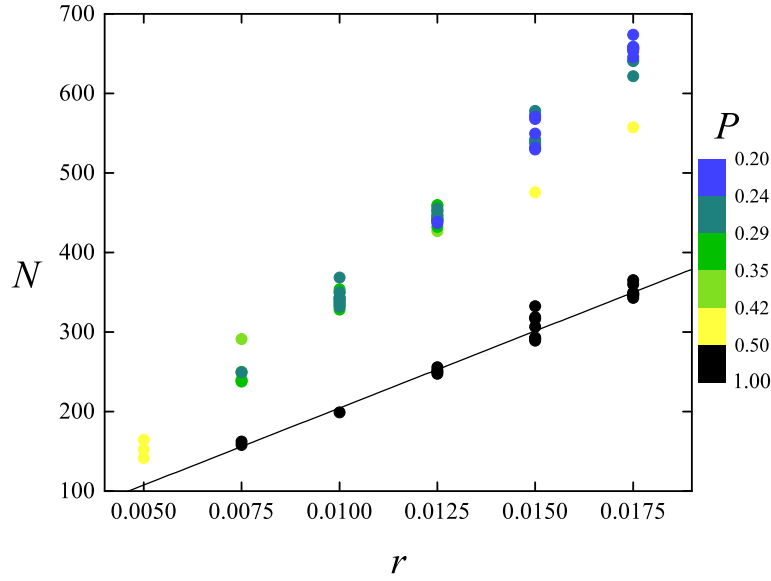
**Figura 5.2:** Módulo cuadrado de la transformada de fourier de la dinámica poblacional en el estacionario. Curvas del tipo  $f^{-\alpha}$  ajustan al espectro con  $\alpha$  igual para valores distintos de la tasa de recuperación del recurso.



**Figura 5.3:** Evolución de la autocorrelación del número de habitantes, obtenido a partir de el espectro de potencias de la Figura 5.2 y el teorema de Wiener-Kinchin.

población consuma menos del recurso, para poder persistir. Como se muestra en la Figura 5.4, esto es lo que efectivamente sucede. En la figura se grafican las poblaciones promedio en el estacionario de las simulaciones para distintas realizaciones y distintos valores de la tasa de recuperación. Como esperábamos, las poblaciones que tienen probabilidad de paso promedio (sobre la población) mayor, son las que tienen menos individuos.

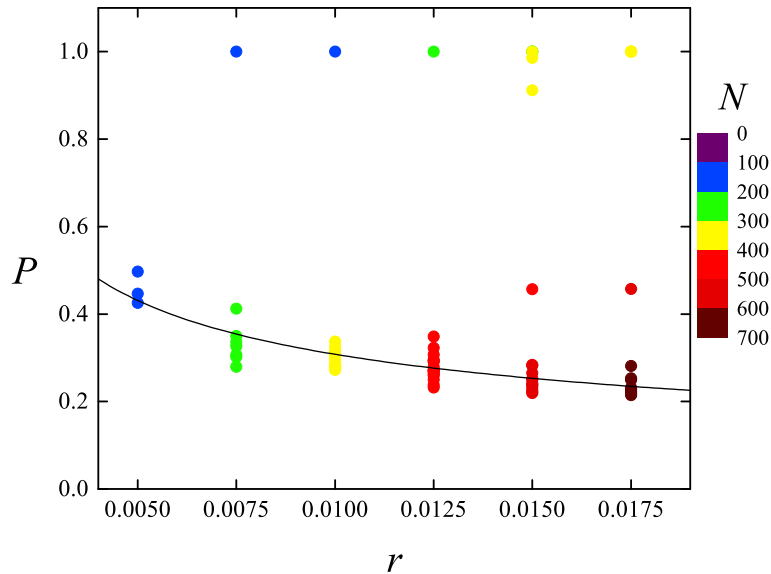
En la Figura 5.5 mostramos las probabilidades de paso medias a las que convergen distintas poblaciones para distintos valores de tasas de recuperación. La rama superior de esta figura se corresponde



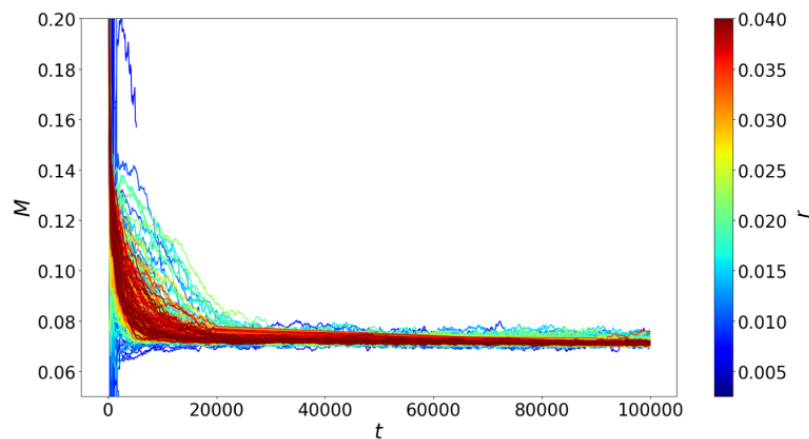
**Figura 5.4:** Poblaciones estacionarias para realizaciones de varias probabilidades de paso. Para probabilidades de paso (energía cinética) grandes, las poblaciones son pequeñas, y viceversa.

con la rama inferior de la Figura 5.4. Sorprendentemente, la rama superior tiene un comportamiento distinto de la inferior: mientras que la primera es constante en  $P = 1$ , la segunda tiene un comportamiento algebraico  $r^\gamma$  con  $\gamma = -1/2$ . Esto sugiere que la rama superior proviene de una situación fuera de equilibrio.

Las dos ramas mostradas en las Figuras 5.4 y 5.5 corresponden a grupos genéticos distintos que surgen de las mismas condiciones iniciales. Por esto decimos que el modelo presenta especiación, a pesar de que las especies no convivan en el sistema. En realidad, la falta de convivencia es de esperar debido al principio de exclusión competitiva [37]. Este afirma que cuando dos especies compiten por el mismo recurso, eventualmente una de las dos extinguirá a la otra. En un sistema espacialmente más complejo (heterogéneo, o estructurado topológicamente), cabría esperar la formación de especies por un mecanismo simpátrico, vale decir por aislamiento geográfico, como suele ocurrir en realidad [15].



**Figura 5.5:** Probabilidades de paso medias de distintas realizaciones y valores de tasas de recuperación. La rama superior es constante, con  $P = 1$ , mientras que la inferior tiene un comportamiento algebraico  $r^{-1/2}$ .



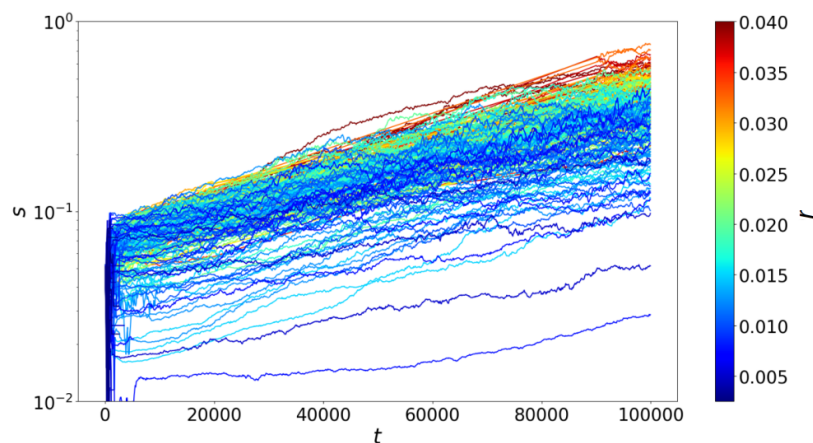
**Figura 5.6:** Evolución temporal de la masa adulta. Vemos que todas las realizaciones y tasas de recuperación del recurso.

La masa adulta, por su parte, tuvo un comportamiento sorprendente e inexplicado. En la Figura 5.6 se muestra la evolución de los valores medios de la masa para distintas realizaciones y velocidades de recuperación. Observamos que las masas adultas siempre convergen al mismo valor.

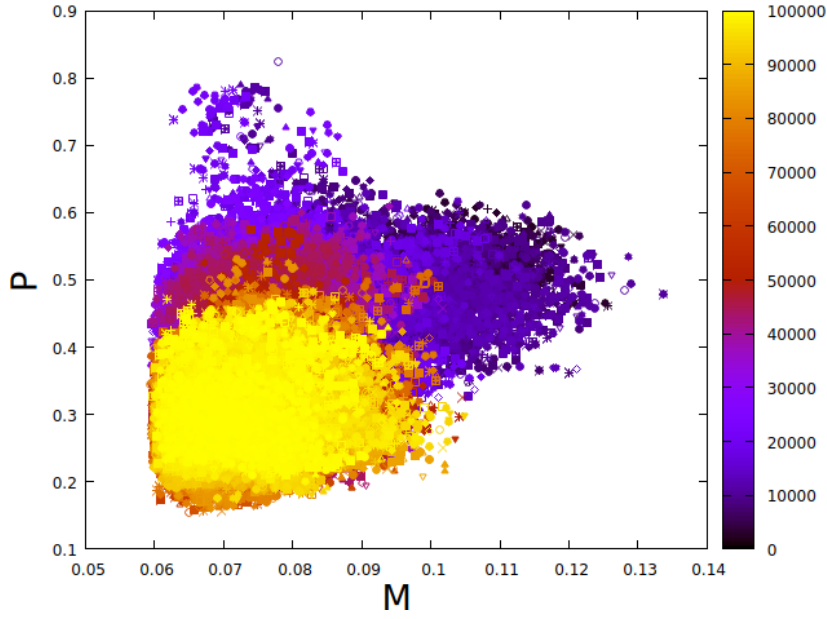
La velocidad de crecimiento, por otra parte, aumentó monótona y exponencialmente en todas las realizaciones. Su evolución, mostrada en la Figura 5.7, sugiere que el costo de aumentar la velocidad de crecimiento es nula, de la misma manera que sucedía en el modelo de la hidra inmortal. Sin embargo, al considerar un costo de crecimiento mayor (multiplicando el modelo usado por un factor mayor que uno), observamos que la velocidad de crecimiento no aumentaba monótonamente, indicando que simplemente es despreciable en la manera que está definida.

También podemos destacar que tanto la velocidad de crecimiento como la masa, y también el número de individuos en la población, muestran una transición bastante abrupta entre el período transitorio y el estacionario donde tienen parámetros aproximadamente constantes. Esto es reminiscente del *equilibrio puntuado* [30], un fenómeno conocido en la biología evolutiva donde los cambios en las especies ocurren en intervalos cortos y permanecen en estacionarios largos, con transiciones abruptas entre ambos.

Para concluir con la caracterización genética del modelo, mostramos la evolución del plano  $(M, P)$



**Figura 5.7:** Evolución temporal de la velocidad de recuperación. Vemos que todas las realizaciones y tasas de recuperación crecen exponencialmente (notar la escala logarítmica) con la misma tasa, sugiriendo que el aumento es ilimitado en esta región. Verificamos esta hipótesis penalizando al crecimiento con un factor de proporcionalidad mayor a uno.



**Figura 5.8:** Evolución temporal del plano  $(M, P)$  del espacio genético. Se observa una tendencia general hacia valores pequeños de los genes hasta ciertos valores restringidos. Las cotas inferiores halladas para los genes son mucho menores a las que se observan realmente.

del espacio genético en la Figura 5.8. Cada punto representa un individuo y los colores indican la ventana temporal. Para no sobrecargar la imagen, elegimos una simulación iniciada con un solo individuo. Se observa una tendencia general de la población a genes más pequeños y más lentos, con claras restricciones para masas menores que 0.06 y probabilidades de paso menores que 0.15. A pesar de que buscamos explicaciones para estos límites, las cotas calculadas resultaron ser muy inferiores a las observadas.

### 5.2.2. Probabilidad de muerte

En el modelo de la hidra, los nacimientos ocurren luego de un proceso bastante complejo. Deben pasar  $c_R = 100$  pasos durante los cuales la densidad de energía  $\varepsilon = E/m$  es mayor o igual al umbral de reproducción,  $\varepsilon_U = 0.8$ . Las muertes, por otra parte, requieren de una condición mucho más simple. A continuación, haremos un intento por calcular su probabilidad de ocurrir. Con esta veremos también que calcular la probabilidad de nacimientos es mucho más difícil.

En primer lugar, nos tomaremos el trabajo de caracterizar mejor la probabilidad de muerte en general, para cumplir con lo prometido en la introducción de este capítulo. Pensando en la historia de vida de un solo individuo, consideremos la probabilidad  $P_D(t)$  de que el mismo haya muerto antes de un tiempo  $t$ . Esta probabilidad es la acumulada de la distribución de probabilidad de tiempos de muerte,  $p_D(t)$ . La distribución de tiempos de muerte se puede pensar como un proceso de Poisson ya que morir a tiempo  $t$  requiere de no morir hasta tiempo  $t$  y en ese momento, encontrarse con un evento letal. Si la tasa de eventos letales, que en general no es constante, es  $\rho_D(t)$ , podemos relacionar  $P_D$ ,  $p_D$  y  $\rho_D$  según:

$$\begin{aligned} p_D(t) &= \rho_D(t) \exp \left( - \int_0^t \rho_D(t') dt' \right), \\ &= - \frac{d}{dt} \left( \exp \left( - \int_0^t \rho_D(t') dt' \right) \right), \end{aligned} \quad (5.1)$$

con lo cual, utilizando  $dP_D(t)/dt = p_D(t)$  y la condición de contorno  $P_D(0) = 0$ ,

$$\begin{aligned}\frac{dP_D(t)}{dt} &= -\frac{d}{dt} \left( \exp \left( -\int_0^t \rho_D(t') dt' \right) \right), \\ \ln(1 - P_D(t)) &= \int_0^t \rho_D(t') dt', \\ \Rightarrow \rho_D(t) &= \frac{p_D(t)}{1 - P_D(t)}.\end{aligned}\tag{5.2}$$

La Ecuación (5.2) relaciona la probabilidad acumulada de muertes, medible de las simulaciones, con la tasa de eventos letales y la distribución de probabilidad de muertes, obtenible de la acumulada, con lo cual podemos obtener cualquiera de las otras dos. Como la muerte es un proceso de Poisson, la última línea de esta ecuación es distributiva en el sentido de que, si la muerte se puede producir por distintos tipos de eventos independientes  $i$ , cada uno de ellos tendrá asociada una tasa de muerte a la probabilidad de ese tipo de eventos,

$$\begin{aligned}\rho_D^i(t) &= \frac{p_D^i(t)}{1 - P_D^i(t)}, \\ \rho_D(t) &= \sum_i \rho_D^i(t).\end{aligned}\tag{5.3}$$

Para implementar la muerte por edad en las simulaciones (uno de los cambios respecto del modelo de la hidra inmortal), usamos la Ecuación (5.3) suponiendo que hay una probabilidad  $P_D^{edad}(t)$  logística que da una vida media de  $\lambda = 1000$  con tasa de cambio del orden de la inversa del tiempo medio de vida, es decir:

$$\begin{aligned}P_D^{edad}(t) &= \frac{1}{1 + \exp(-2(t - \lambda)/\lambda)}, \\ \Rightarrow p_D^{edad}(t) &= \frac{2}{\lambda} \frac{\exp(-2(t - \lambda)/\lambda)}{(1 + \exp(-2(t - \lambda)/\lambda))^2}, \\ \Rightarrow \rho_D^{edad}(t) &= \frac{2}{\lambda} \frac{1}{1 + \exp(-2(t - \lambda)/\lambda)}.\end{aligned}\tag{5.4}$$

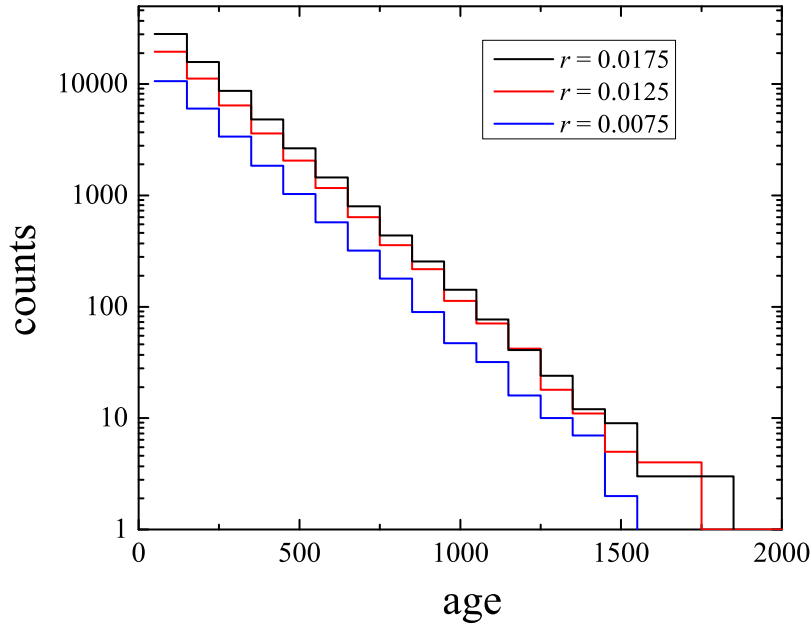
Así, la distribución  $p_D^{edad}$  es la que se implementa en las simulaciones para determinar las muertes por edad, e implica una tasa  $\rho_D^{edad}$ . Esta última es importante porque es la que se puede obtener de la distribución de edades de la población, como mostraremos a continuación.

En primer lugar, consideremos una población caracterizada por una distribución de edades  $z(t)$ . El tiempo al cual esta distribución se encuentra no importa porque estudiaremos su comportamiento en el estacionario, por lo cual  $t$  denota la edad, no el tiempo. Si la tasa de eventos letales (por individuo) a esa edad es  $\rho_D(t)$ , la distribución evoluciona de acuerdo a:

$$\begin{aligned}z(t + dt) &= z(t) - z(t)\rho_D(t)dt, \\ \frac{dz(t)}{dt} &= -z(t)\rho_D(t), \\ \Rightarrow \rho_D(t) &= -\frac{d \ln(z(t))}{dt}.\end{aligned}\tag{5.5}$$

Midiendo la distribución de edades de la población en las simulaciones y usando la última línea de la Ecuación (5.5), podemos obtener la tasa de muerte total.

En la Figura 5.9 se muestra la distribución  $z(t)$ . Es una exponencial casi perfecta (notar la escala logarítmica del eje  $y$ ), con lo cual  $\rho_D(t) = 0.055$  es una constante. Esto indica que la edad media de la población es de aproximadamente  $\tau = 181$ , mucho menor que  $\lambda = 1000$  disponible por la muerte por envejecimiento. Esto sugiere que, en seres vivos reales, el envejecimiento que ocurre por errores de copia debe ser lo suficientemente lento como para que se lleguen a dar suficientes reproducciones y



**Figura 5.9:** Distribución de edades  $z(t)$  para tres tasas de recuperación distintas. Las distribuciones son exponenciales (rectas en escala logarítmica con la misma pendiente) con la misma tasa de eventos letales.

la especie sobreviva.

### 5.2.3. Comportamiento altruista

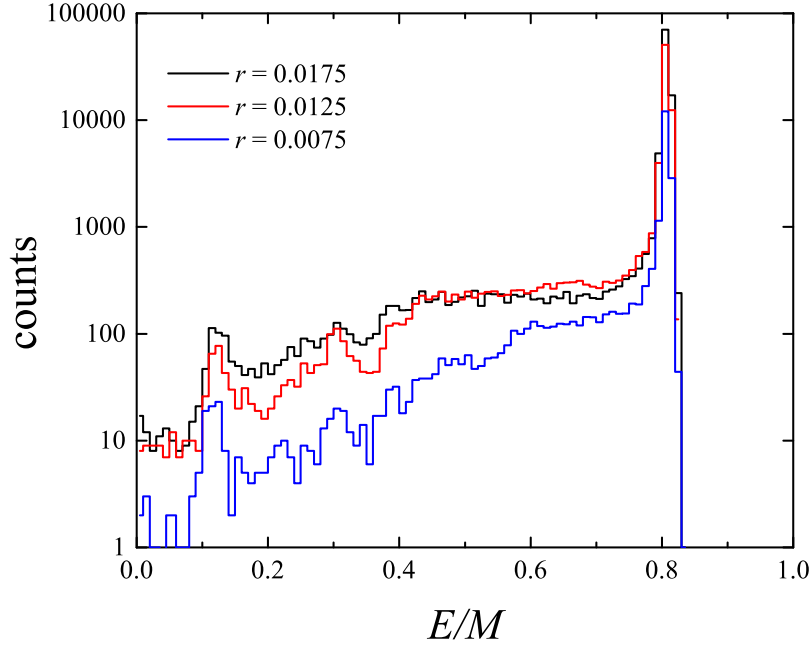
El hecho de que la vida media de la población se acorte desde los  $\lambda = 1000$  disponibles en principio hasta los  $\tau = 180$  observados en la distribución de población expresa una especie de comportamiento altruista por parte de la población. En lugar de vivir todo lo que pueden y producir más descendencia, los individuos del sistema estacionario tienen una vida media relativamente corta. Esta es suficiente para una o dos reproducciones, a diferencia de las nueve o diez que podrían tener. Es un comportamiento reminiscente de lo que ocurre en países con alta calidad de vida [55], donde la población disminuye la cantidad de hijos por mejorar su propia calidad de vida. Además, es contraintuitivo respecto a lo que se esperaría maximizando la *fitness*, entendida en el sentido biológico estricto como el número de descendientes.

Otro comportamiento altruista observado en las simulaciones, mucho más obvio que el del acortamiento de la vida, se expresa en la distribución de energía de la población. En la Figura 5.10 mostramos distribuciones de densidad de energía medidas sobre la población estacionaria. La gran mayoría de la población (notar la escala logarítmica del eje vertical) tiene una densidad de energía de  $\varepsilon = 0.8$ , que define el umbral de energía  $\varepsilon_U$  necesario para la reproducción.

Esto es sorprendente ya que la retribución  $\varphi$  (definida en el modelo de la hidra inmortal, Ecuación (4.4)) implica que los individuos deben consumir el recurso local hasta llenar su reserva. Sin embargo, la evolución conjunta de la población y el recurso encuentra una forma de evitar esta restricción, dejando a los individuos con la mínima energía necesaria para reproducirse sin quitarle a los otros más de lo imprescindible para que el sistema continúe existiendo.

Si revisamos la Figura 4.11, con la distribución de energía del modelo de la hidra inmortal, vemos que aquella no tiene el mismo comportamiento. El máximo se encuentra en  $\varepsilon = 1$ , como cabría de esperar. Esto sugiere que la posibilidad de evitar la restricción de máximo consumo está relacionada con la independencia entre la probabilidad de paso y la masa. Efectivamente, levantar esta restricción hace que cada uno de los genes pueda optimizarse exclusivamente en términos de su función.





**Figura 5.10:** Distribuciones de densidad de energía medida en la población estacionaria. La gran mayoría de la población no tiene densidad de energía 1, como se esperaría con

A continuación, estudiaremos el sistema a través de su dinámica media, tratando de lograr un mejor entendimiento de su funcionamiento.

#### 5.2.4. Campo medio

Estudiando la evolución de valores medios de propiedades del modelo podemos concretar las ideas discutidas sobre el consumo del recurso y su relación con la dinámica poblacional. Para empezar, consideremos el consumo de un individuo adulto. Estudiamos el consumo de un individuo adulto porque, como vemos en la Figura 5.7, la velocidad de crecimiento aumenta sin restricciones, por lo cual podemos despreciar el período de crecimiento en la mayoría de la población. En un intervalo de tiempo  $t$ , un individuo adulto de genotipo  $\mathbf{G} = (M, s, P)$  (y fenotipo  $\mathbf{F} = \mathbf{G}$  porque es adulto), que se mueve  $n$  veces consume:

$$C(t, n, \mathbf{F}) = \mu M^{3/4} t + \kappa M m, \quad (5.6)$$

con probabilidad binomial,

$$P(n|t, \mathbf{F}) = \binom{t}{m} (1 - P)^{t-m} P^m. \quad (5.7)$$

Así, el consumo medio por unidad de tiempo es:

$$C(\mathbf{F}) = \mu M^{3/4} + \kappa M P. \quad (5.8)$$

Supondremos que los individuos consumen la misma cantidad de recurso que gastan, así pueden mantener su energía constante a lo largo de sus vidas. Si la dinámica energética de un individuo fuese negativa, sus descendientes empezarían sus vidas con menos energía y, como son similares, también tendrían una dinámica negativa. Eventualmente, el linaje se extinguiría. Si la dinámica fuese positiva, eventualmente siempre se acumularía energía suficiente como para reproducirse y la dinámica del linaje sería exponencial como se mostró en el modelo de la hidra inmortal, en la Sección 4.1. Como el recurso no es inagotable, sabemos que esta dinámica es insostenible y por lo tanto no puede ser parte

del estado estacionario. Con este argumento, podemos calcular la cantidad total de recurso consumido en un sitio, que denotaremos  $a(\mathbf{x})$ . Si  $\mathcal{P}(\mathbf{x})$  es el subconjunto de la población  $\mathcal{P}$  que se encuentra en el sitio  $\mathbf{x}$ , el consumo total en un tiempo  $\Delta t$  en ese sitio es:

$$a(\mathbf{x}) = \min \left( \sum_{c \in \mathcal{P}(\mathbf{x})} C(\mathbf{F}_c) \Delta t, f(\mathbf{x}) - K\beta \right). \quad (5.9)$$

En el límite  $\Delta t \rightarrow 0$ , el consumo del recurso siempre es igual al primer argumento del mínimo. Esto nos permite escribir:

$$\begin{aligned} \frac{d\langle f \rangle_{\mathbf{x}}}{dt} &= \frac{1}{L^2} \sum_{\mathbf{x}} \left[ r f(\mathbf{x}) \left( 1 - \frac{f(\mathbf{x})}{K} \right) - \frac{a(\mathbf{x})}{dt} \right], \\ &= \frac{1}{L^2} \sum_{\mathbf{x}} r f(\mathbf{x}) \left( 1 - \frac{f(\mathbf{x})}{K} \right) - \frac{N}{L^2} \frac{1}{N} \sum_{\mathbf{x}} \sum_{c \in \mathcal{P}(\mathbf{x})} C(\mathbf{F}_c), \\ &= r \left[ \langle f \rangle_{\mathbf{x}} \left( 1 - \frac{\langle f \rangle_{\mathbf{x}}}{K} \right) - \frac{\text{var}_{\mathbf{x}}(f)}{K} \right] - \frac{N}{L^2} \langle C \rangle_{\mathcal{P}}, \end{aligned} \quad (5.10)$$

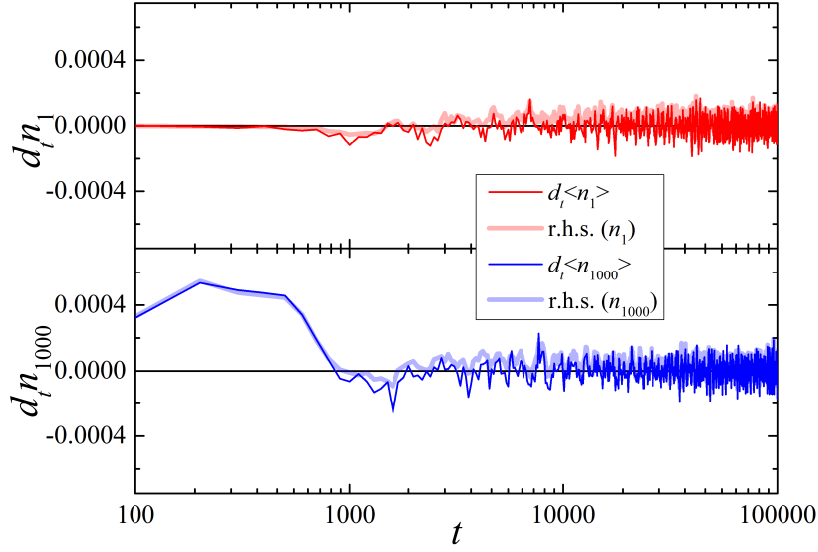
donde el subíndice  $\mathbf{x}$  indica que los valores medios se toman sobre la distribución espacial. Por su parte, el subíndice  $\mathcal{P}$  indica que el promedio es sobre la población. Notar que suponemos que la tasa de recuperación y la capacidad de carga son constantes en el dominio del recurso, pero esto no necesariamente es así. Sin embargo, dado que son características constantes en cada sitio, deben ser estadísticamente independientes al recurso. El resultado general se obtiene reemplazando  $r$  por  $\langle r \rangle_{\mathbf{x}}$  (y lo mismo para  $K$ ).

Una forma de simplificar la Ecuación (5.10) es normalizando el recurso. Haciendo el cambio de variables  $n = f/K$  se obtiene:

$$\frac{d\langle n \rangle_{\mathbf{x}}}{dt} = r [\langle n \rangle_{\mathbf{x}} (1 - \langle n \rangle_{\mathbf{x}}) - \text{var}_{\mathbf{x}}(n)] - \frac{N}{L^2} \frac{\langle C \rangle_{\mathcal{P}}}{K}. \quad (5.11)$$

En la Figura 5.11 se muestran los lados izquierdo y derecho de la Ecuación (5.11) calculados a partir de datos obtenidos de simulaciones. La curva superior corresponde a una simulación con un solo miembro inicial, mientras que la inferior empieza con 1000 individuos. La primera ( $n_1$ ) se mantiene siempre en equilibrio mientras que la segunda ( $n_{1000}$ ) tiene un período transitorio inicial fuera del equilibrio. Sin embargo, en ambas simulaciones, el lado izquierdo y derecho de la Ecuación (5.11) coinciden casi a la perfección, validando el resultado de campo medio. El hecho de que funcione fuera del equilibrio sugiere que las fluctuaciones energéticas que le provocan la muerte a los individuos son mucho más rápidas que las escalas de tiempo típicas de la vida de los mismos. Si no fuera así, la suposición de que el consumo energético es igual a la ingesta ocurriría más a menudo en la vida de un individuo, y el lado derecho de la ecuación tendría una expresión que combinase los argumentos del mínimo en la Ecuación (5.9) de acuerdo a su probabilidad.

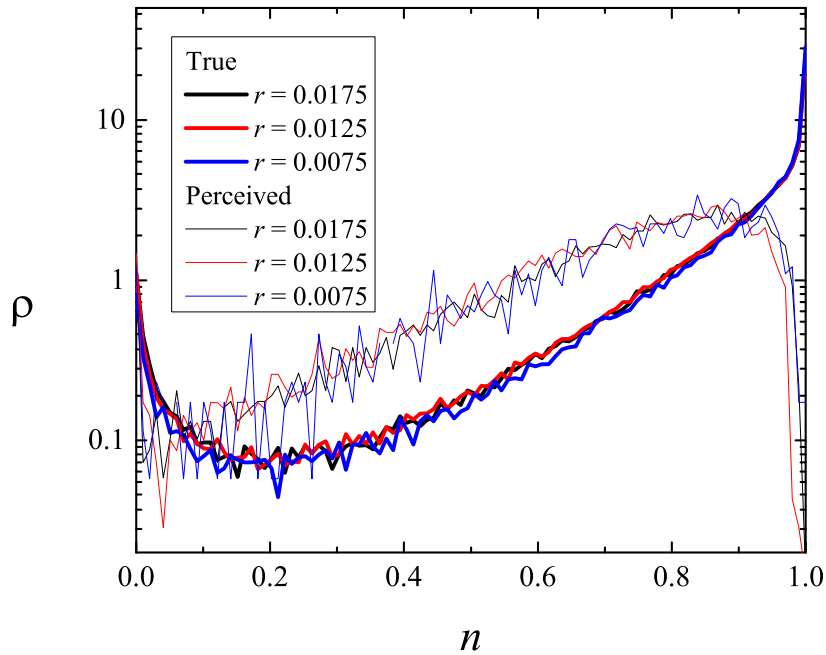
La ecuación (5.11) sugiere una interpretación del sistema población-recurso como un sistema depredador-presa, respectivamente. La interacción entre ambos se da por el término proporcional a  $N$ . Su relación con el recurso “real”  $f$  está dado por la Ecuación (5.9) y la normalización  $n = f/K$ . Para describir completamente el sistema, se requiere una segunda ecuación que describa a los depredadores. En la Sección 4.1 se propusieron algunos de estas ecuaciones para el modelo de la hidra inmortal, pero ninguna resultó particularmente satisfactoria. Hallar la dinámica poblacional es un tema notoriamente difícil, por lo cual lo dejaremos por el momento y continuaremos con el estudio del recurso.



**Figura 5.11:** Comparación de los lados derecho e izquierdo de la Ecuación (5.11) para una simulación que empieza con 1 individuo ( $n_1$ ) y una que empieza con 1000 ( $n_{1000}$ ). La primera se mantiene en equilibrio a lo largo de su evolución mientras que la segunda presenta un periodo transitorio inicial. En ambos casos, los lados derecho e izquierdo coinciden, validando la ecuación.

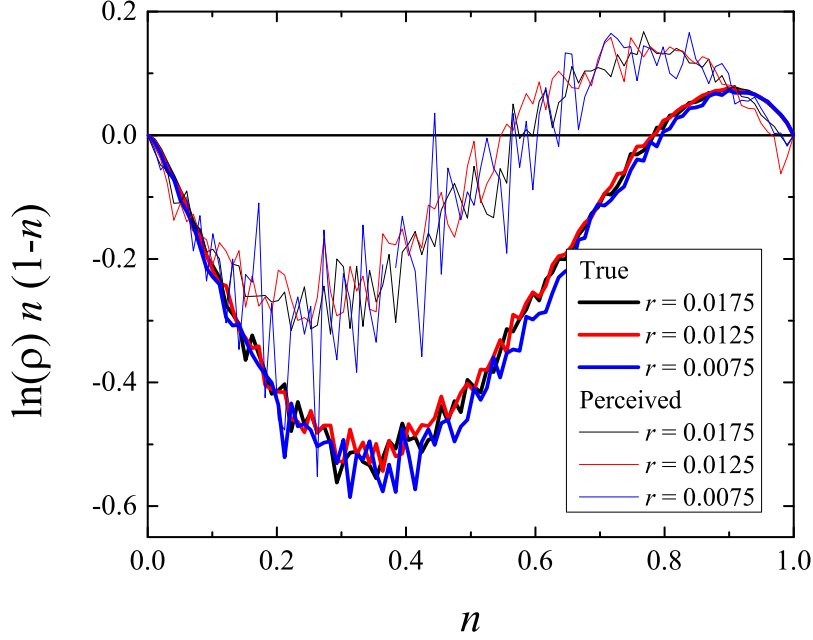
### 5.3. Dinámica del recurso

La Figura 5.12 muestra, en trazos gruesos (“True”), la distribución de recurso  $\rho(n)$  en la grilla para tres tasas de recuperación. En trazos finos (“Perceived”) muestra la distribución restringida a sitios ocupados  $\rho(n|o)$ , es decir la distribución que percibe la población realmente. En ambos casos, las distribuciones son indistinguibles entre valores distintos de la tasa de recuperación. Las distribuciones se obtuvieron por histogramas tomados sobre tiempos en el estacionario de la evolución de población, para realizaciones con 1000 individuos iniciales.



**Figura 5.12:** Distribución del recurso sobre toda la grilla (“True”) y restringido a los sitios ocupados (“Perceived”). En ambos casos las distribuciones son indistinguibles entre valores de la tasa de recuperación, dentro del error estadístico.

Las distribuciones de la Figura 5.12 presentan picos marcados en los extremos  $n = 0$  y  $n = 1$ . Para caracterizarlas mejor, mostramos en la Figura 5.13 las distribuciones “regularizadas”, transformadas según  $\ln(\rho)n(1-n)$ . Podemos observar así que ambas distribuciones tienen forma cúbica o sinusoidal, de amplitudes levemente distintas y desfasadas.



**Figura 5.13:** Distribución del recurso mostradas en la Figura 5.12 transformadas según  $\ln(\rho)n(1-n)$  para evitar posibles divergencias en los extremos del rango de  $n$ .

Podemos relacionar la distribución percibida y la real de la siguiente manera. En primer lugar, la distribución del recurso se puede descomponer según sitios ocupados ( $o$ ) y desocupados ( $\neg o$ ),

$$\rho(n) = \rho(n|o)\rho(o) + \rho(n|\neg o)\rho(\neg o). \quad (5.12)$$

Esto permite separar la dinámica temporal de la distribución en dos partes:

$$\begin{aligned} \frac{d\rho(n)}{dt} &= \frac{d\rho(n)}{dn} \frac{dn}{dt}, \\ &= \frac{d\rho(n|o)}{dn} \rho(o) \frac{dn}{dt} \Big|_o + \frac{d\rho(n|\neg o)}{dn} \rho(\neg o) \frac{dn}{dt} \Big|_{\neg o}. \end{aligned} \quad (5.13)$$

La distribución de sitios ocupados,  $\rho(o)$  es simplemente la densidad espacial de la población,  $N/L^2$ , y la de desocupados  $\rho(\neg o)$  es su complemento. Con estos valores en la Ecuación (5.12) podemos despejar  $\rho(n|\neg o)$  en términos de cosas conocidas o medibles e insertarlo en la Ecuación (5.13). Así, esta toma la forma:

$$\begin{aligned} \frac{d\rho(n)}{dt} &= \frac{d\rho(n|o)}{dn} \rho(o) \frac{dn}{dt} \Big|_o + \left\{ \frac{d\rho(n)}{dn} - \frac{d\rho(n|o)}{dn} \rho(o) \right\} \frac{dn}{dt} \Big|_{\neg o}, \\ &= \frac{d\rho(n|o)}{dn} \rho(o) \left\{ \frac{dn}{dt} \Big|_o - \frac{dn}{dt} \Big|_{\neg o} \right\} + \frac{d\rho(n)}{dn} \frac{dn}{dt} \Big|_{\neg o}. \end{aligned} \quad (5.14)$$

Para completar la relación entre la distribución verdadera y la percibida en equilibrio, faltan las expresiones para  $dn/dt$  en los casos ocupado y desocupado. El segundo es simplemente la dinámica logística del recurso. Para el primero, suponemos que hay solo un individuo en cada sitio ocupado

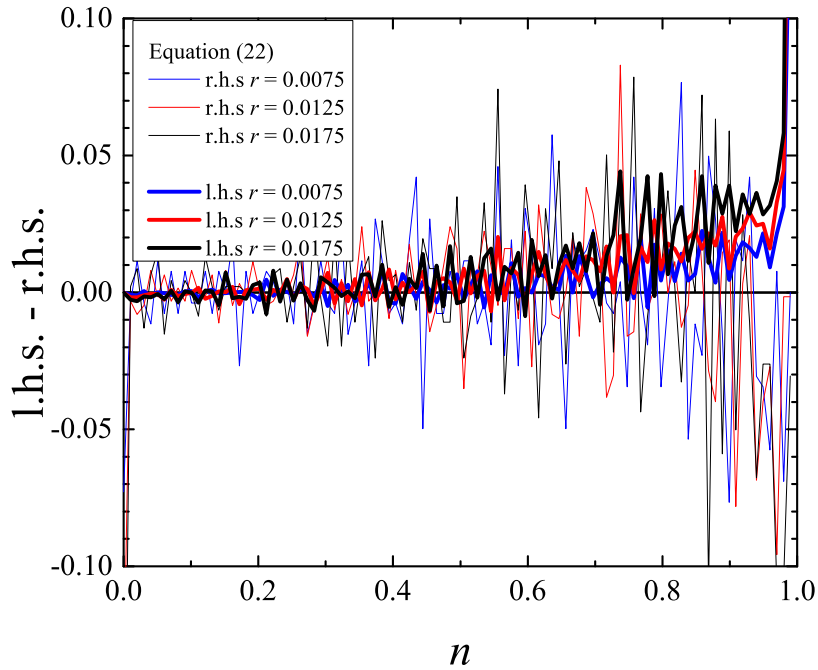
(debido a la baja densidad poblacional) con lo cual las expresiones para  $dn/dt$  son:

$$\begin{aligned} \left. \frac{dn}{dt} \right|_{-o} &= rn(1-n), \\ \left. \frac{dn}{dt} \right|_o &= rn(1-n) - \frac{\langle C \rangle}{K}. \end{aligned} \quad (5.15)$$

Ahora, suponiendo  $d\rho(n)/dt = 0$  y reemplazando todo en la Ecuación (5.13), obtenemos:

$$rn(1-n) \frac{d\rho(n)}{dn} = \frac{N}{L^2} \frac{\langle C \rangle}{K} \frac{d\rho(n|o)}{dn}. \quad (5.16)$$

En la Figura 5.14 graficamos los dos miembros de la Ecuación (5.16). Notamos que las curvas tienen un comportamiento parecido hasta aproximadamente  $n = 0.7$  pero después se separan, divergiendo en direcciones opuestas. Atribuimos esta separación a que el recurso se encuentra fuera de equilibrio en estas regiones: si tenemos en cuenta las Figuras 5.12 y 5.13, la población prácticamente no puede percibir recurso en abundancia  $n \approx 1$ . Esto se debe a que la dinámica de crecimiento del recurso es cercana a cero en esta región y por lo tanto el crecimiento es mucho más lento que la recuperación. Matemáticamente, la condición de equilibrio de la Ecuación (5.13) igualada a cero necesita, para tomar derivada, que la velocidad a la que el consumo de la población “mueve” diferenciales de recurso hacia valores menores converja a cero a la misma velocidad que el crecimiento del mismo. Como el consumo de la población es discreto en la implementación simulada, el extremo  $n \approx 0$  no cumple con esta condición porque  $dn/dt \approx 0$ . En otras palabras, la evolución del recurso no es una función continuamente diferenciable en esta región debido al rápido consumo (respecto al crecimiento). En el extremo inferior, donde  $n \approx 0$ , la velocidad de crecimiento es igual de lenta, pero el efecto no es observable porque la población no sobrevive en estas celdas, y los individuos que las ocupan mueren demasiado rápido como para tener efecto.



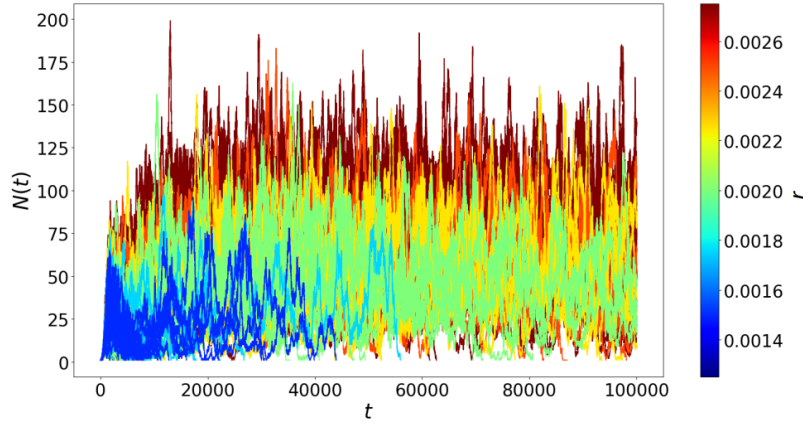
**Figura 5.14:** Lados izquierdo (l.h.s.) y derecho (r.h.s.) de la Ecuación (5.16). Ambos lados parecen coincidir hasta  $n \approx 0.7$ , donde los efectos del consumo discreto de la población se vuelven apreciables respecto a la velocidad de recuperación del recurso. En esta región la hipótesis de cambio continuo en el recurso, necesaria para plantear la Ecuación (5.13) e igualarla a cero, se vuelve inválida. En la región  $n \approx 0$  el efecto no es observable ya que los individuos que llegan a estas celdas mueren demasiado rápido como para afectar las curvas.

## 5.4. Algunas modificaciones

### Evas

Hasta aquí, salvo que lo aclarásemos, trabajamos con simulaciones que comenzaban con 1000 individuos uniformemente distribuidos en el espacio genético. Hicimos esto principalmente para que la muestra genética inicial fuera lo suficientemente grande como para mostrar los efectos estocásticos en la simulación. Efectivamente, hallamos dos grupos genéticos mutuamente excluyentes de las Figuras 5.4 y 5.5 en estas condiciones.

Sin embargo, actualmente se cree que la vida se originó de un solo individuo [56, 57]. Para probar si nuestro modelo es compatible con esa idea, hicimos simulaciones que comenzaban con un solo individuo. Los datos de  $dn_1/dt$  de la Figura 5.11 se obtuvieron de tal simulación. Observamos (y la misma figura lo demuestra) que estas simulaciones se mantienen mucho más cerca del equilibrio a lo largo de su evolución comparado con las de población inicial mayor. Esto también se expresa en la Figura 5.15. En esta observamos poblaciones que sobreviven para valores de  $r$  mucho menores que las halladas en las simulaciones con poblaciones iniciales grandes mostradas en la Figura 5.1. También vimos que, para las simulaciones con un solo individuo inicial, las probabilidades de paso de la simulación siempre convergían al valor menor entre las dos ramas de la Figura 5.5, confirmando que la superior proviene de una situación fuera del equilibrio. Los resultados de este experimento sugieren que no solo es posible que la vida haya comenzado con un solo individuo, sino que pareciera ser más probable.



**Figura 5.15:** Población en función del tiempo comparando simulaciones que comienzan con un solo individuo. Observamos que las poblaciones sobreviven a valores mucho más bajos que los que requieren las simulaciones con mil individuos iniciales mostrados en la Figura 5.1.

### Movimiento inteligente

Otra modificación que hicimos al modelo fue vincular la probabilidad de movimiento con la reserva de energía. En *Dugesia tigrina*, un pequeño invertebrado de agua dulce con un sistema nervioso rudimentario, se detectó un comportamiento guiado por la necesidad de alimentarse [58]: cuando el organismo había comido recientemente se movía poco, y cuando tenía hambre, se movía mucho para encontrar más alimento. Este comportamiento es esperado, y ocurre en seres vivos mucho más complejos (incluso en humanos [59]), pero también podría converger a la solución inversa para ahorrar la poca energía que queda cuando falta la comida. Implementamos esto cambiando la probabilidad de paso instantánea en función de la masa instantánea, por una probabilidad de paso modulada por la cantidad de energía,  $p = a + bE$ , con  $a$  y  $b$  genes. Efectivamente, observamos que, casi siempre,

$b$  tomaba valores negativos y bastante uniformes en la población, emulando el comportamiento de *Dugesia tigrina*. Por otra parte el valor de  $a$ , que determina el umbral de energía al que comienza el movimiento (dado  $b$ ), estaba bastante uniformemente distribuido.

Estas modificaciones en el modelo de la hidra pueden no ser las más precisas desde un punto de vista de la similitud con la realidad. Sin embargo, creemos que capturan bien los comportamientos cualitativos de algunas propiedades biológicas. El hecho de que los resultados de las simulaciones coincidan con lo observado en la naturaleza (dentro del margen que permite una descripción cualitativa) nos deja conformes sobre la posibilidad de estudiar modelos más complejos y detallados con este método.





## Capítulo 6

# Discusión y conclusiones

*“Der anfang ist die ende und der ende ist der anfang.”*

— Adam

A lo largo de la historia, las propias teorías evolutivas han estado sujetas a la selección natural en manos de la ciencia. La primera teoría, con los seres humanos puestos en un pedestal por su creador, fue de las que más tiempo duró, aunque no por acertada. Poco a poco, personas muy valientes comenzaron a cuestionar esta idea, basándose en evidencia hasta entonces descorrelacionada. La figura de Darwin se yergue como central y culminante en este desarrollo, al cambiar las especies estáticas por un escenario en el que mutan aleatoriamente, y en el cual las mutaciones más favorables tienden a producir más descendencia. Llamó a este principio la supervivencia de los más aptos, y con él sentó las bases de la teoría evolutiva moderna. Con el auge de la computación y las simulaciones numéricas, muchos modelos de evolución y selección fueron explorados para sustentar la teoría de manera cualitativa y cuantitativa en un marco formal y matemático. Bak y Sneppen, por ejemplo, presentaron un modelo que aumentaba la “aptitud” global a partir de interacciones entre pares y el reemplazo de los menos aptos, pero sin definir qué es la aptitud exactamente [53]. Abramson y Zanette [60, 61], por su parte, han mostrado cómo las extinciones obedecen a una estadística realista sin necesidad de optimizar una *fitness*. Por otro lado, la teoría se extendió más allá de su alcance biológico, y actualmente algunos problemas de optimización utilizan un algoritmo inspirado en la selección natural, proponiendo soluciones, variándolas al azar y seleccionando las mejores (más aptas) según el problema a optimizar. Es el caso, por ejemplo, de mi compañero de curso Julián Neñer, en cuya tesis de Maestría [62] los agentes optimizan su desempeño en transacciones económicas. En todos estos casos, la aptitud se conoce como *fitness* y, en el ámbito de la biología, se suele asociar con la probabilidad de producir descendencia fértil, pero esta es una definición operativa.

En esta tesis hemos propuesto un modelo para describir la evolución por selección natural de los seres vivos que no está basado en una *fitness*. El modelo plantea un balance entre la necesidad de cada individuo de conseguir energía para sobrevivir, compitiendo por un recurso compartido, y el costo energético asociado a conseguirla. Esto convierte el problema de optimización (de la *fitness*) en una restricción: la energía de los individuos debe ser positiva para estar vivos. Cada individuo tiene un fenotipo que determina cómo se obtiene y consume la energía. De generación en generación, transmite su genotipo, que determina el fenotipo. Implementamos este modelo con simulaciones numéricas, proponiendo varios casos particulares y estudiando los resultados de cada uno, buscando rasgos generales y propiedades emergentes.

En primer lugar estudiamos un modelo basado en redes booleanas donde cada estado fenotípico está dado por una cadena de  $L$  bits y tiene asociado un costo, una retribución máxima (limitada por

la disponibilidad de recurso) y un estado consecutivo asignados al azar. Cada genotipo, que coincide con la cadena de bits que se utiliza como estado inicial, tiene asociado un fenotipo particular en el cual el individuo se reproduce, insertando otro individuo en la población. En primera instancia consideramos una situación en la cual no se producían errores en la copia del genotipo durante la reproducción. Caracterizamos la evolución del sistema para distintas dinámicas de recuperación del recurso. Observamos que un recurso ilimitado producía distintos tipos de crecimiento poblacional en diversos genotipos: crecimiento exponencial, lineal y extinciones. Además, recurso limitado pero de recuperación inmediata (representando un flujo constante de energía, como los fotones del Sol) producían una población constante. Finalmente, dinámicas de recuperación continua resultaban en la extinción de la población debido a la independencia de la tasa de reproducción y el recurso. En una segunda etapa incorporamos errores de copia mediante un parámetro de similitud que representa cuán probable es que la descendencia sea igual a su progenitor. Variando esta calidad de copia de simulación en simulación, hallamos que uno de los grupos genéticos maximizaba su población para el caso de copia perfecta mientras que los linajes tienen poblaciones similares para el caso de copia totalmente aleatoria. Un hecho interesante es que el linaje que maximizaba su población no era el de mayor tasa de reproducción en el crecimiento exponencial, indicando que el sistema no maximiza *fitness* en el sentido convencional. Además, la población total se mantuvo aproximadamente constante en todos los casos. Al dejar que la similitud mute como el genotipo (en función de la propia similitud), hallamos que se producía una transición de calidad de copia baja (con una población heterogénea) a calidad de copia alta (y población de un solo linaje). Finalmente, incorporamos un parámetro de herencia, que determina cuánto de su energía acumulada los progenitores ceden a sus descendientes. Observamos que el linaje con más éxito (de mayor población) sin la entrega de energía disminuía su población y luego se extinguía al entregarle más de un 40 % de su energía a la descendencia, siendo gradualmente reemplazada por la siguiente en orden de población decreciente.

Para poder estudiar mejor los efectos de la similitud genética entre una generación y la siguiente planteamos una simplificación de nuestra versión del modelo de redes booleanas. En esta implementación interpretamos el costo y el beneficio máximo como variables del fenotipo en sí, y no como propiedades asociadas a este. Completando el fenotipo de cada individuo dispusimos una fracción de energía heredada y una tasa de reproducción. El genotipo siempre se definió como igual al fenotipo, suponiendo individuos adultos. Caracterizamos la velocidad de crecimiento de la población en función de la cantidad de energía entregada a la descendencia, hallando que esta es máxima para la mínima herencia. El resto de los genes tenían distribuciones uniformes y hallamos que sus valores medios siempre presentaban evoluciones algebraicas con el tiempo. Luego permitimos que esta entrega de energía sea un gen variable, evolucionando como los otros. Esto permitió a la población crecer mucho más rápido que lo que podía cuando el gen no mutaba, sugiriendo que los grados de libertad del genotipo juegan un papel en su capacidad de evolución. Observamos cómo disminuye la entropía de la distribución genética al tender los valores medios de los genes a sus valores asintóticos a costa de un consumo cada vez más eficiente (por individuo pero no por población total) del recurso. Luego estudiamos la variación de los resultados en función de la similitud genética, que definía el ancho de la distribución de los genes de la descendencia en torno al genotipo de los progenitores. Observamos que la población aumenta con la similitud, pero presentaba un cambio de fase para calidad de copia perfecta debido al condicionamiento del sistema a sus condiciones iniciales. Finalmente, propusimos una manera de extender este modelo a que los genes tengan correlación entre sí y argumentamos cómo esto eventualmente evoluciona hacia una red neuronal.

En el Capítulo 4 repetimos los resultados del modelo de la hidra inmortal, el modelo de selección natural basado en energía que estudiamos durante mi tesis de licenciatura. Planteamos un modelo inspirado en seres vivos macroscópicos, biológicamente inmortales y de reproducción asexual y un

recurso distribuido espacialmente. Debido a reglas de reproducción más complejas en este modelo (un indicador de calidad del entorno, medido a partir de la cantidad de recurso hallado), la dinámica del recurso podía ser logística, a diferencia de los casos de los capítulos anteriores. Cada uno de los individuos cuenta con un tamaño corporal que define su metabolismo y velocidad de movimiento (mediante un tamaño de paso discreto) por la grilla donde está distribuido el recurso. Comparamos su dinámica poblacional con la de modelos conocidos (crecimiento exponencial de Malthus, modelo logístico y Lotka-Volterra). Cada uno de estos modelos presentó más características comunes con las simulaciones que el anterior, pero ninguno las describió exhaustivamente. Mostramos algunas propiedades que distinguen a nuestro modelo, tales como la probabilidad de supervivencia en función de la velocidad de recuperación y la saturación de la cantidad del recurso. En el primer caso, la probabilidad resultó siempre creciente con la velocidad y decreciente en el tiempo, como es de esperarse, pero presentando “escalones” de probabilidad de supervivencia estable. En el segundo caso, para saturaciones grandes del recurso, observamos que en algunos casos, para saturaciones crecientes, la probabilidad disminuía. También estudiamos los períodos de oscilación hallados en la población, que mostraron reglas algebraicas en función de la tasa de recuperación (con exponente  $-1/2$ ) y la capacidad de carga (exponente  $1/2$ ). Finalmente, la dinámica y distribución genética mostraron siempre comportamientos parecidos para minimizar el tamaño de paso en la grilla. Contraintuitivamente, hallamos que los individuos se volvían más grandes luego de períodos con menos recurso disponible y más pequeños para períodos de recurso abundante.

Finalmente, estudiamos una modificación del modelo de la hidra inmortal. Consideramos una probabilidad de paso en lugar de un tamaño del mismo y observamos que las simulaciones convergían a poblaciones rápidas, de pocos habitantes, o lentas y de muchos. Interpretamos este resultado en el sentido de que el modelo presentaba especiación, si bien no coexistencia. Describimos la evolución de los genes para distintas tasas de recuperación de recurso y medimos la distribución del mismo. En este contexto, hallamos comportamientos altruistas rudimentarios a partir de reglas egoistas. Por ejemplo, los individuos fueron programados para ingerir el recurso en cantidad suficiente como para llenar su reserva de energía (o lo que estuviera disponible). De todos modos, encontramos que el sistema evolucionó hacia una situación donde los individuos ingerían lo necesario para reproducirse, pero no lo suficiente como para llenar sus reservas. El análisis de la dinámica del modelo de la hidra constituye la principal contribución de esta tesis.

Para este modelo obtuvimos ecuaciones de campo medio para la dinámica del recurso y encontramos que la derivada de su valor medio servía de parámetro para caracterizar la distancia del equilibrio de las simulaciones. Finalmente, con pequeñas modificaciones a las simulaciones, observamos que estas se mantenían más cerca del equilibrio (y las poblaciones sobrevivían más consistentemente) con un solo individuo inicial en lugar de varios con mayor diversidad genética. También incorporamos movimiento “inteligente” de manera muy simple, inspirado en un ser vivo real, y observamos que el comportamiento de las simulaciones convergió al mismo que este organismo presenta.

Quedó pendiente relacionar, de manera general, la evolución de los genes con el costo y la retribución asintóticos asociados, lo cual tiene el potencial para transformar el problema en términos de restricciones en uno de optimización. Tampoco alcanzamos a estudiar cómo cambian los comportamientos de estos sistemas al considerar un vector de energías, como sería necesario si se tuviera en cuenta la necesidad utilizar distintos nutrientes. Encontramos que, a medida que los modelos de consumo y retribución se vuelven más similares a organismos reales, los resultados de las simulaciones también. Concluimos que basar la evolución en una restricción, como es mantener la energía positiva, es una metodología que tiene el potencial para abarcar e incluso reemplazar los métodos de *fitness* usados actualmente.



## Apéndice A

# Pseudocódigo

La población está ordenada como una lista doblemente anidada donde cada una de sus componentes es un miembro de la población. Cada miembro tiene los siguientes elementos:

- Genotipo.
- Fenotipo.
- Energía.
- Lugar para almacenar un miembro en gestación (descendencia).
- Variables booleanas para el control de la gestación.

El genotipo, el fenotipo y la energía siguen la dinámica descrita en el Capítulo 1 y de acuerdo al modelo particular. El lugar para almacenar un miembro en gestación se implementó como un puntero a otro miembro que no puede almacenar a un tercero. Las variables booleanas son dos que se encienden (1) solo al instante de inicio y finalización de la gestación. La variable de inicio de la gestación se actualiza solo cuando el lugar para almacenar un miembro en gestación está vacío y el de fin de la gestación, solo cuando el lugar se encuentra vacío.

```

Data: tiempo_simulación
Result: Correr simulación completa
crear población;
crear recurso;
while tiempo < tiempo_simulación y tamaño(población) > 0 do
    for miembros en población do
        actualizar fenotipo de miembro;
        actualizar gestación de miembro;
        actualizar energía de miembro según recurso;
        if miembro está vivo then
            if miembro empezó a gestar then
                agregar miembro_en_gestación en miembro;
            end
            if miembro terminó de gestar then
                nuevo_miembro = miembro_en_gestación;
                quitar miembro_en_gestación de miembro;
                agregar nuevo_miembro a población;
            end
        else
            quitar miembro de población;
        end
    end
end

```

**Algorithm 1:** Pseudocódigo del algoritmo general de optimización de los individuos en una población. Para cada modelo en particular, el algoritmo de actualización del fenotipo, gestación y energía (y recurso) será distinto, produciendo otras dinámicas de nacimientos y muertes.

# Bibliografía

- [1] Harari, Y. From animals into gods: A brief history of humankind. CreateSpace Independent Publishing Platform, 2012.
- [2] Encyclopedia Britannica. The history of biology. URL <https://www.britannica.com/science/biology/The-history-of-biology#ref48827>, last accessed 2nd Feb. 2021.
- [3] Aristóteles. Metafísica. 1311-1321 (Versión Andrónico de Rodas).
- [4] Brack, A. Chapter 10.4 - Clay minerals and the origin of life. En: F. Bergaya, G. Lagaly (eds.) Handbook of Clay Science, tomo 5 de *Developments in Clay Science*, págs. 507–521. Elsevier, 2013.
- [5] Aristotle (Translation by D’Arcy Wentworth Thompson). The history of animals. 350 BCE.
- [6] al Batriq, Y. I. Kitab Al-Hayawan (Book of animals). 850 (First mention).
- [7] Savage-Smith, E. Attitudes toward dissection in medieval Islam. *Journal of the History of Medicine and Allied Sciences*, **50**, 67–110, 1995.
- [8] Vesalius, A. De humani corporis fabrica (On the fabric of the human body). School of Medicine, Padua, 1543.
- [9] Otten, W. Medieval scholasticism: Past, present, and future. *Nederlands Archief voor Kerkgeschiedenis / Dutch Review of Church History*, **81** (3), 275–289, 2001.
- [10] Pomata, G., Siraisi, N. G. Historia: Empiricism and erudition in early modern Europe. MIT Press, 2005.
- [11] Magner, L. N. A history of the life sciences, revised and expanded. CRC Press, 2002.
- [12] Popham, A. E. The drawings of Leonardo da Vinci. Reynal & Hitchcock, New York, 1973.
- [13] Ribatti, D. An historical note on the cell theory. *Experimental cell research*, **364** (1), 1–4, 2018.
- [14] Wolpert, L. Evolution of the cell theory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **349** (1329), 227–233, 1995.
- [15] Darwin, C. On the origin of species. John Murray, London, 1859.
- [16] Borisov, S., Podberezskaya, N. X-ray diffraction analysis: A brief history and achievements of the first century. *Journal of Structural Chemistry*, **53** (1), 2012.
- [17] Kirz, J., Jacobsen, C. The history and future of X-ray microscopy. En: J Phys Conf Ser, tomo 186, págs. 1–11. 2009.
- [18] Malthus, T. R. An essay on the principle of population. J. Johnson, London, 1872.

- [19] Malthus, P. Notice sur la loi que la population poursuit dans son accroissement. *Corresp. Math. Phys.*, **10**, 113–121.
- [20] Kingsland, S. Alfred J. Lotka and the origins of theoretical population ecology. *Proceedings of the National Academy of Sciences*, **112** (31), 9493–9495, 2015.
- [21] Azar, A. T., Vaidyanathan, S. Advances in chaos theory and intelligent control, tomo 337. Springer, 2016.
- [22] Robertson, R., Combs, A. Chaos theory in psychology and the life sciences. Psychology Press, 2014.
- [23] Carroll, S. The big picture: on the origins of life, meaning, and the universe itself. Penguin, 2017.
- [24] Luisi, P. L. About various definitions of life. *Origins of Life and Evolution of the Biosphere*, **28** (4), 613–622, 1998.
- [25] Spafford, E. H. Computer viruses as artificial life. *Artificial life*, **1** (3), 249–265, 1994.
- [26] Anzalone, A. V., Koblan, L. W., Liu, D. R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology*, **38** (7), 824–844, 2020.
- [27] Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., *et al.* Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell stem cell*, **13** (6), 659–662, 2013.
- [28] Munnink, B. B. O., Nieuwenhuijse, D. F., Stein, M., O’Toole, Á., Haverkate, M., Mollers, M., *et al.* Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nature medicine*, **26** (9), 1405–1410, 2020.
- [29] Dorigo, M., Birattari, M., Stutzle, T. Ant colony optimization. *IEEE computational intelligence magazine*, **1** (4), 28–39, 2006.
- [30] Eldredge, N., Gould, S. J. Punctuated equilibrium prevails. *Nature*, **322**, 211–212, 1988.
- [31] Kaneko, K. Life: An introduction to complex systems biology. Springer, 2006.
- [32] Ornes, S. Core concept: How nonequilibrium thermodynamics speaks to the mystery of life. *Proceedings of the National Academy of Sciences*, **114** (3), 423–424, 2017.
- [33] Smith, J. M. Optimization theory in evolution. *Annual review of ecology and systematics*, **9** (1), 31–56, 1978.
- [34] Schrödinger, E. What is Life? With Mind and matter and Autobiographical sketches. Cambridge University Press, 1992.
- [35] Schneider, E. D., Kay, J. J. Life as a manifestation of the second law of thermodynamics. *Mathematical and computer modelling*, **19** (6-8), 25–48, 1994.
- [36] Martyushev, L. M. Life defined in terms of entropy production: 20th century physics meets 21st century biology. *BioEssays*, **42** (12), 2070115, 2020.
- [37] Hardin, G. The competitive exclusion principle. *Science*, **131**, 1292–1297, 1960.
- [38] Kauffman, S. A. The origins of order: Self-organization and selection in evolution. Oxford University Press, 1993.
- [39] Pate, R. R. The evolving definition of physical fitness. *Quest*, **40** (3), 174–179, 1988.



- [40] Pasteur, L. Spontaneous generation. *The lancet*, **109**, 332, 1877.
- [41] Roll-Hansen, N. Revisiting the Pouchet–Pasteur controversy over spontaneous generation: understanding experimental method. *History and Philosophy of the Life Sciences*, **40**, 68, 2018.
- [42] Eigen, M., Schuster, P. The Hypercycle. A principle of natural self-organisation. Springer, New York, 1979.
- [43] Nei, M. Stochastic errors in dna evolution and molecular phylogeny. *Prog Clin Biol Res*, **218**, 133–147, 1986.
- [44] Walsh, C. Posttranslational modification of proteins: Expanding nature’s inventory. Roberts and Company Publishers, 2006.
- [45] Abadi, N. Mecanismos de selección natural en modelos de sistemas biológicos, Diciembre 2019. Tesis de Licenciatura en Física. Instituto Balseiro, Universidad Nacional de Cuyo.
- [46] Murray, J. D. Mathematical biology: I. An introduction. Springer Science & Business Media, 2007.
- [47] Murray, J. D. Mathematical biology. Springer, 2002.
- [48] Kleiber, M., *et al.* Body size and metabolism. *Hilgardia*, **6** (11), 315–353, 1932.
- [49] Zhou, Z., Zou, X. Stable periodic solutions in a discrete periodic logistic equation. *Applied Mathematics Letters*, **16** (2), 165–171, 2003.
- [50] Stephens, P. A., Sutherland, W. J., Freckleton, R. P. What is the Allee effect? *Oikos*, **87**, 185–190, 1999.
- [51] Twitchett, R. J. The Lilliput effect in the aftermath of the end-Permian extinction event. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **252** (1-2), 132–144, 2007.
- [52] Abadi, N., Abramson, G. An energy-based natural selection model, 2020. Enviado, preprint en: <https://arxiv.org/abs/2011.06945>.
- [53] Bak, P., Sneppen, K. Punctuated equilibrium and criticality in a simple model of evolution. *Physical Review Letters*, **71**, 4083, 1993.
- [54] Karlin, S., Taylor, H. M. A first course in stochastic processes. 2<sup>a</sup> ed<sup>ón</sup>. Academic Press, 2012.
- [55] Westendorp, R. G., Kirkwood, T. B. Human longevity at the cost of reproductive success. *Nature*, **396** (6713), 743–746, 1998.
- [56] Ponting, C. P., Russell, R. B. Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all  $\beta$ -trefoil proteins. *Journal of molecular biology*, **302** (5), 1041–1047, 2000.
- [57] Briones, C., Fernández Soto, A., Bermúdez de Castro, J. M. Orígenes. El universo, la vida, los humanos. Crítica, 2015.
- [58] Sheiman, I., Zubina, E., Kreshchenko, N. Regulation of the feeding behavior of the planarian *Dugesia (Girardia) tigrina*. *Journal of Evolutionary Biochemistry and Physiology*, **38**, 414–418, 07 2002.

- 
- [59] Lowden, A., Moreno, C., Holmbäck, U., Lennernäs, M., Tucker, P. Eating and shift work - effects on habits, metabolism, and performance. *Scandinavian Journal of Work, Environment & Health*, **36**, 150–162, 2010.
- [60] Abramson, G., Zanette, D. Statistics of extinction and survival in lotka-volterra systems. *Physical Review E*, **57**, 4572–4577, 1998.
- [61] Abramson, G. Ecological model of extinctions. *Physical Review E*, **55**, 785–788, 1997.
- [62] Neñer, J. Redes neuronales para la búsqueda de estrategias óptimas en problemas de econofísica, Febrero 2021. Tesis de Maestría en Ciencias Físicas. Instituto Balseiro, Universidad Nacional de Cuyo.

# Publicaciones asociadas

1. N. Abadi, G. Abramson, An energy-based natural selection model, **Physica A.**, Nov. 2020 (entregada). <https://arxiv.org/abs/2011.06945>



# Agradecimientos

Antes que nada, agradezco a la gente a quien está dedicado este trabajo, no saben cuánto los aprecio. Gracias a todos los educadores que hicieron un esfuerzo por enseñarme un poco de lo que sabían, tanto en el ámbito académico como no. Aprendí muchísimo. Gracias a Seba Bouzat, quien participó y aportó a muchas discusiones sobre el modelo de la hidra inmortal, que forma parte de mi tesis de licenciatura pero sienta las bases de este trabajo. Gracias a la CNEA y UNCuyo por la experiencia maravillosa de estos últimos cuatro años. Y gracias a todo el resto.

